Essays in public finance and macroeconomics

DISSERTATION
of the University of St. Gallen,
School of Management,
Economics, Law, Social Sciences
and International Affairs
to obtain the title of
Doctor of Philosophy in Economics and Finance

submitted by

Thomas Davoine
from
Neuchâtel

Approved on the application of

Prof. Dr. Christian Keuschnigg
and

Prof. Dr. Volker Grossmann

Dissertation no. 4255

epubli GmbH, Berlin 2013
The University of St. Gallen, School of Management, Economics, Law, Social Sciences and International Affairs hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St. Gallen, November 15, 2013

The President:

Prof. Dr. Thomas Bieger
# Contents

Foreword ................................................. i
Executive Summary (English) ...................................... ii
Executive Summary (German) ....................................... iv
Introduction .................................................. vi

1 Flexicurity and Job Reallocation .......................... 1
    1 Introduction ........................................... 2
    2 The Model ............................................. 7
    3 Optimal Flexicurity .................................... 13
    4 Policy Reform .......................................... 21
    5 Comparative Statics of Optimal Policy .................. 30
    6 Conclusions ............................................ 35

A Proof details ............................................. 37
B Calibration details ........................................ 40

2 Labor supply and skill premiums ......................... 47
    1 Introduction .......................................... 48
    2 Literature overview .................................. 50
    3 Model ................................................. 54
    4 Analysis ............................................... 56
    5 Numerical illustration with representative agents .... 67
    6 Historical patterns on labor supply in the cross-section 70
    7 Concluding remarks .................................... 72

A Proofs ................................................... 75

3 Time constraints and consumer behavior ............... 89
    1 Introduction .......................................... 90
    2 Literature overview and discussion ...................... 92
    3 Simple counterfactual experiment ....................... 98
Foreword

This document is the thesis that I submit for the obtention of the PhD in Economics and Finance degree from the University of St.Gallen. I started the PhD programme in September 2008 and submitted the thesis 4 years later, in September 2012. After discussions of the submission in January 2013, revisions have been applied until September 2013. Much of the research contained in the thesis was performed while I was a research assistant at the Chair of Prof. C. Keuschnigg between April 2009 and December 2011.

The research contributions that I present in this thesis cover four different topics of public finance and macroeconomics. Together with Prof. C.Keuschnigg, we investigate the role of flexicurity policies on job reallocation, unemployment and welfare. I then investigate labor supply decisions when skill differences between workers change over time. I later consider the role of time constraints in consumption on economic outcomes and differences in savings behavior between households. I finally analyze the effect of Eurobonds on public debt financing and trade.

None of the research delivers ready-to-use policy recommendations but most of it is relevant for the design of public policy. We did not come to definitive conclusions on flexicurity, but we analyzed in what capacity they may help reduce unemployment. It appears that the cost and efficiency of active labor market policies, such as job search assistance to the unemployed, are key determinants of the value of flexicurity. I did not consider any policy instrument in the analysis of labor supply decisions, nor in the analysis of time constraints on consumption. Both however help understand differences in decisions, in particular why rich households save at a higher rate. A correct understanding of this behavior is necessary for the design of redistributive yet efficient policy, for instance. Eurobonds is a topic of current policy debate. I contribute directly to the debate and offer conclusions on the desirability of Eurobonds when we neglect unemployment and its financing. The latter is a limitation but the analysis is a first step.

Compared to the research proposal that I presented in September 2010, there has been one change of topic. I replaced the empirical analysis of the effect of progressive taxation on labor supply by the theoretical analysis of Eurobonds and public debt sustainability. The decision was motivated by practical considerations unrelated to science. I believe that the quality of the research has not been diminished by this change. Indeed, both topics involved new beginnings, compared to the two other topics covered by the research. Ability to perform empirical research still remains an important goal of mine.

I am very grateful to Prof. Christian Keuschnigg for his valuable support and guidance. I also thank the other members of the thesis committee, Prof. Michael Lechner, Prof. Volker Grossmann and Prof. Martin Kolmar. Friends and relatives further provided help in many ways. I am grateful to all of them, including my family, Jochen, Philip, Evelyn, Niklaus, Mirela, Alexander as well as Andreas, Andrei, Anthony, Charlotte, Conny, Daniel, Darjusch, Florian, Giovanni, Johannes, Lukas, Martin, Petra, Philipp, Rudi, Sandra, Uwe and Yazid.
Executive Summary (English)

My PhD thesis contains four essays on distinct topics of public finance and macroeconomics.

The first deals with flexicurity, which has been associated with reduced unemployment in some countries. Research was carried with my advisor, Prof. C. Keuschnigg.

Employment protection reduce inflows into unemployment, but also depresses job creation because firms can not close jobs easily if something goes wrong. The idea of flexicurity is to let firms fire freely to encourage job creation and compensate fired workers with generous unemployment benefits and assistance in getting new jobs.

Building a model, we prove that the introduction of unemployment insurance increases welfare. So do job protection if skills are sector specific and job search assistance if it is not costly. Numerical illustrations show that all policy instruments used in flexicurity help, but flexicurity itself may not be optimal.

The second topic deals with differences in labor supply between high- and low-skilled households. High-wage US households worked fewer hours in 1920, but more in 1980. I analyze the role of exogenous changes to a skill premium in a standard neoclassical labor supply model. I show that high-skilled households provide more labor than low-skilled if the premium increases long enough, and vice-versa. With unexpected labor tax increases, every household will reduce its labor supply. High-skilled households may work more than low-skilled before and after a tax increase, even if both reduce working hours.

The starting point of the third topic is that consumption needs time. In a standard neoclassical model, I remove the assumption that labor supply is constant and I assume that time is needed for consumption. Households decide how to split time between consumption and production. I show that as technology improves, more goods can be produced and time needs to be shifted from production to consumption in order to enjoy these extra goods. Consumption expenditures and output will grow slower than productivity.

Workers need to coordinate labor supply while consumption can be performed alone. I assume that high and low earning households need to agree on the amount of labor to supply. High earning households want to shift more time to consumption than low earning households as technology improves, because they can afford more consumption. However, the need to coordinate prevents them from shifting as much time as they want. They end up saving at a higher rate than low income households. This theoretical prediction is consistent with empirical observation.

The fourth topic is Eurobonds. Investors who lend to governments are not protected against repayment defaults, as governments define the law. Investors charge an interest rate premium on public debt, which increases with fear of defaults. Eurobonds could help resolve public debt problems in some Eurozone countries and are currently debated. I use a model where the premium depends not only on current debt, but also on future
expected income and expenses of government. Countries with same debt-to-output ratios pay then different premiums. I show that the premium increases in a convex fashion under certain circumstances: the more investors doubt future repayment capacities of a country, the higher the premium.

Numerical simulations show that Eurobonds are successful in improving debt sustainability but come at a welfare cost over the long run, a surprising result similar to the transfer paradox.
Executive Summary (Deutsch)


Kündigungsschutz reduziert Zufluss in die Arbeitslosigkeit, sondern auch die Schaffung von Arbeitsplätzen, weil Firmen Arbeitsplätze schwer schließen können, wenn etwas schlecht geht. Die Flexicurity Idee ist für Firmen Arbeitsplätze frei zu schließen, um Arbeitsplätzen Schaffung zu fördern, hoch Arbeitslosengeld zu geben und Unterstützung Jobsuche zu geben.


Numerische Simulationen zeigen, dass Eurobonds Schuldenträgerfähigkeit verbessern, aber mit langfristigem Wohlfahrtskosten, ein überraschendes Ergebnis ähnlich der Transfer Paradox.
Introduction

My thesis consists of four contributions to economic theory on different topics, distantly related. I start with a general motivation and then discuss each topic in turn. The introduction is meant to be accessible to an audience of non-specialists. It is written with as little technicality as possible and without references to the literature, which are left for the detailed investigations presented in subsequent chapters. This introduction is not meant to be a substitution for the standard introductions in each chapter. Experts interested in some or all of the detailed investigations can skip this introduction.

1 General motivation

The four contributions to economic theory that I provide follow a standard pattern in the profession. Theories start with a simple model, making simplifying assumptions in order to obtain theoretical predictions and gain understanding of real economic processes. Models are then developed by replacing those assumptions which deviate most from reality with more realistic assumptions and modified to correct theoretical predictions which are inconsistent with empirical evidence. In this spirit, I focus on three different inconsistencies of neoclassical economic theory, namely perfect labor markets, homogeneous consumer behavior and complete debt contracts. None of these are novel research topics but in all of them remain open questions.

Early macroeconomic theories assumed perfect labor markets, labor demand being equal to labor supply. In reality, a part of unemployment is involuntary: some unemployed workers would like to work but can not. In my first contribution, joint work with Prof. C. Keuschnigg, I analyze a policy design which is sometimes considered to be promising to reduce unemployment, the so-called flexicurity.

Another assumption of early macroeconomics is that all households are the same or that they only differ by age. There are many other differences, in reality. In my second contribution, I consider the impact of differences in skill on labor supply, when the skill difference is constant or varies over time. In my third contribution, I consider the impact of skill differences in a model where time is needed for production and consumption. The main goal is to explain why high- and low-skilled households save at a different rate.

In a large part of public finance, debt contracts are assumed to be complete. Contracts
themselves and institutions guarantee that loans will be repaid to their full extent. In reality, governments can default and investors will charge different interest rates to reflect the risk of default. In my fourth contribution, I allow for incomplete debt contracts and investigate the effect of common bonds in a currency area, such as Eurobonds.

With one exception, there are no strong links between the four topics of my thesis, for reasons unrelated to science. The exception are the second and third contributions. In both cases, households differ by a given skill level, which they keep forever. In the second contribution, I focus on the impact of changing skill premiums and compare labor supply of the high- and low-skilled. In the third contribution, I keep the skill premium constant. Insights gained in the second contribution are useful for the third contribution.

In the following sections I present more in details each of the four contributions.

2 Imperfect labor markets: flexicurity and job reallocation

Europe has been struggling with high unemployment for over three decades. Flexicurity is an example of policy predating economic science. It consists of generous unemployment benefits, low employment protection and active labor market policies, such as job search assistance or sanctions for low search efforts. Governments in Northern Europe, with traditionally high levels of unemployment insurance and flexible markets, increased the importance of active labor market programmes in the 1990’s. Flexicurity has been associated with reduced unemployment, as illustrated for Denmark in figure 1: compared to the European average, the unemployment rate declined more rapidly after generalization of active labor market policies.

![Unemployment rate (%)](image)

Figure 1: Unemployment rates in Europe and Denmark, 1970-2008.

The economic literature has since then been considering flexicurity in a more systematic way. Several models consider constrained flexicurity, in the sense that unemployment insurance and active labor market policies can vary but employment protection is main-
tained at a low value. The model that I use is the first with no such restrictions. One can then compare flexicurity to other policy designs.

One key feature of flexicurity is the low employment protection. Several countries have a generous unemployment insurance and provide active labor market policies, but few have the level of flexibility that is offered to firms under flexicurity. Theory shows, and empirical evidence generally supports, that job protection reduces inflows into unemployment, as expected, but also inflows out of unemployment. Indeed, firms are reluctant to hire new workers if it is difficult to separate from them under negative circumstances. The net effect remains ambiguous.

The idea of flexicurity is to allow firms to fire workers without any restrictions and compensate unemployed workers with a high level of unemployment insurance and assistance to find a new job.

The essay contained in my thesis is joint work with Prof. C. Keuschnigg. Further work, not included in the thesis, provide an extended numerical illustration.

We build a static model where workers are free to choose between two sectors. For simplicity, we consider only one sector with risky jobs. Choosing the risky sector is rewarded with a higher wage but the risk of losing the job: in case of a negative productivity shock, the job is terminated and the worker fired. If the shock is only mildly negative and employment protection (in the form of a firing tax) is high, the firm will prefer to keep the worker. If fired, the worker then perceives unemployment benefits and looks for a new job. We assume that new jobs can only be found in the safe sector to focus on the effect of active labor market policies, which involve retraining programmes. Depending on the magnitude of these policies, the unemployed worker will find more or less easily a job in the safe sector. Unemployment insurance also plays a role as generous benefits discourage search efforts.

Theoretical analysis shows that welfare increases if one introduces unemployment insurance financed by a flat tax on labor income, starting from no government. The conclusion is not a surprise, given the absence of private markets for insurance. Job protection further increases welfare if sectors are very skill specific, so that wages in the risky sector are sufficiently large. This conforms with intuition, as the welfare loss of a fired employee is larger if its wages in the risky sector were higher. We also show that the welfare benefit of active labor market policies depends on the utility gains of these policies compared to their cost. These policies are more likely to deliver welfare benefits if large unemployment benefits are provided. The reason is that every time an unemployed finds a new job the heavy tax burden of unemployment insurance is reduced, so it is worth investing more into job search assistance.

Numerical illustrations show that each policy instrument involved in flexicurity packages delivers welfare benefits. Under standard economic conditions, pure flexicurity is not the optimal welfare policy: unemployment insurance is slightly below the usual level found in welfare states, job search assistance is large but job protection is significant.
3 Heterogeneous consumer behavior: the effect of skill premium variations on labor supply

One reason for inequality are skill differences, which lead to wage differences. Who from low- and high-skilled households work more has varied through time and still varies between countries. Figure 2 shows that high-wage earners worked less in the US early in the 20th century and more towards its end. Quantitative studies are able to reproduce late 20th century labor supply patterns in the US but it is hard to isolate the reasons for the change.

![Graph showing average daily work hours, selected wage deciles, US](source: selection from Costa (2000))

The textbook neoclassical model has ambiguous predictions. At low wage levels, the substitution effect dominates so labor supply increases with wages. At high wage levels, the income effect dominates and labor supply decreases with wages. When wage differences come from an exogenous constant skill premium and the skill inequality is the same as the initial wealth inequality, high- and low-skilled households provide the same labor supply.

I consider a simple extension, allowing the skill premium to vary over time. Development of tertiary education is a more recent phenomenon than development of other education levels and may have benefited high-skilled more than other households. The skill premium has increased over the past few decades, especially in the US.

I analytically show that high-skilled households provide more labor supply if the skill premium increases long enough, and vice-versa. Given education systems evolutions, it is consistent with broad historical patterns in the US. The intuition for such an outcome is the presence of a relative intertemporal labor supply effect when the skill premium exogenously varies over time: if high-skilled households know the premium will increase over time, they have an incentive to work less today and more tomorrow, relative to low-skilled households; when variations takes place long enough, they end up working more.
Skill premium variations are not the only reason why the substitution and income effects do not cancel out. Subsistence consumption could be another. Subsistence needs indeed push low-wage earners to provide more labor than dictated by the distaste for effort. With a wage increase, the subsistence pressure is relaxed so that workers will not increase much labor supply. There is a friction on the substitution effect. For high-wage earners, subsistence is not a worry so the income effect is not constrained and could dominate the substitution effect. Another possible influence is the need for time to consume, which I introduce and discuss in section 4.

4 Heterogeneous consumer behavior: the effect of time use in consumption

In the standard life-cycle model of consumer behavior, households save to finance consumption during retirement. Under homothetic preferences, rich and poor households will save at the same rate: earnings and consumption values of the rich at any point in time are simply a multiple of those of the poor. Yet, empirical evidence show that rich households save a larger fraction of their life-time income, as illustrated in figure 3.

![Figure 3: Households savings rate by income quintiles, US, 1984-1989.](source)

Different theories have been developed to account for this and other empirical patterns. Bequest motives explain the evidence for instance, as poor parents have little incentive to leave a bequest to likely richer kids. However, data shows that rich childless households also save at a higher rate, so the bequest motives explanation needs to be completed. Accounting for different empirical patterns at the same time is often challenging. The model I develop provides an explanation for savings rate heterogeneity and labor market evidence, namely secular declines in per capita labor supply. To the best of my knowledge, there is no other model which can explain these facts at the same time.

I take seriously the hypothesis that rich simply lack time to consume. As wealth is very concentrated, the savings decisions of billionaires and other wealthy households
have disproportionate impacts on aggregate outcomes. I start from a neoclassical Ramsey growth model where I drop the assumption of inelastic labor supply and I am more specific on the consumption process. Following the home production and consumption model of Becker, the act of consumption takes time and consumer goods in themselves deliver no utility. A bike alone delivers no utility, unless it is associated with time to make a bike ride. Households thus need to decide how to allocate a fixed time endowment to either production or consumption.

The first theoretical prediction concerns macroeconomic outcomes. If households want to derive benefits from technological progress, they need to shift some time from production to consumption: if time is needed for consumption and technology improves, how to benefit from higher consumption if no extra time is allocated for consumption of these new goods? As a consequence of this time shifting from production to consumption, I formally show that output and consumption expenditures will grow at a slower pace than technological progress over the long run.

The second theoretical prediction is variation of labor supply over time. Shifting time from production to consumption means that labor supply will decline over time, as technology improves. The prediction is broadly consistent with average labor supply patterns over the 20th century in developed countries. For instance, average weekly work hours declined from 60 in 1900 to 40 in 1960 in the US.

Adding a condition on labor supply and skill premium to the model, the third theoretical prediction is that rich save at a higher rate. As in section 3, I consider households separated by a given constant skill premium difference. Observations and firm theory show that workers coordinate their labor supply: workers tend to work during the same hours. It is intuitive that firms where workers do not synchronize their working hours would be priced out of the market, given the extra delays in production. Without explaining where it comes from but to be consistent with evidence and existing theories, I assume that high-skilled and low-skilled households need to coordinate and provide the same labor supply. Households are free to consume alone but they need to coordinate with other households for production.

The combination of labor supply coordination and time need for consumption can explain why rich households save at a higher rate. Their higher earnings give them access to more production. Compared to poor households, they would thus like to shift more time to consumption as technology improves. However, the need to coordinate labor supply prevents them from shifting as much time as they want. They end up earning more and consuming less than in an autarky world. In other words, they save more. I provide analytical proofs for this result.

In the model, high-skilled would prefer to work less than low-skilled households if there were no coordination constraints. The same outcome would be obtained if there was disutility of labor. Both outcomes are consistent with the results of section 3. Indeed, the need of time for consumption introduces a friction in the substitution effect but not on the income effect, giving a negative slope to the labor supply curve. Because some
time is needed for consumption of additional purchases made possible by a higher wage, households would not increase as much labor supply as in a case where no time is needed for consumption. On the other hand, need of time for consumption has no impact on the income effect.

The model is also consistent with other pieces of evidence on consumer behavior, such as the fact that individuals die with positive assets and asymmetric excess sensitivity, the fact that household consumption responds more to unpredicted income decreases than unpredicted income increases. The model also appears suited for research extensions on several fronts.

5 Incomplete debt contracts: Eurobonds

In textbook public finance theory, debt contracts are complete. Creditors always recover their loans and the interest rate they charge, thanks to complete contractual arrangements and enforcement. In reality, legal environment offer different degrees of investor protection. At one extreme, loans to government are not protected at all because governments define the legal environment. The mere existence of a market for public debt has occupied the early sovereign debt literature. At another extreme, loans between private agents of the same country offer full protection to the creditor.

What is the influence of incomplete debt contracts? Figure 4 illustrates the case of public debt. In 2010, governments of European countries had access to capital markets at roughly the same conditions. In early 2012, interest rates diverge markedly: several countries needed to pay more than 10% in interest rates while other paid less than in 2010. One interesting point is that interest rates rose in spite of the fact that no governments defaulted on their debt repayments, at the time. The mere possibility of defaults influence interest rates.

![Figure 4: Interest rates 10 year government bonds, various EU countries, 2010-2012.](image)

Common bonds giving the same low interest rate to all countries in the Eurozone -
Eurobonds - could be a way to reduce payment burdens and help restore financial sustainability in highly indebted countries. I compare Eurobonds with standard reforms, including fiscal consolidation.

I use a standard overlapping generations model where labor is supplied endogenously, depending on taxes and wages, in a multi-country setting. The same composite good can be traded between countries. The main innovation is the modeling of the interest rate premium on public debt. This premium depends not only on the current level of debt but also on the expected future income and expenses of the government. This forward-looking element can explain why some countries have access to cheap debt in spite of a high debt-to-output ratio, such as Japan, while other countries need to pay more for a lower ratio, including many European countries.

I borrow and adjust the premium definition from recent studies on sovereign debt. Investors lend to governments up to their repayment capacity, the expected net fiscal surpluses. Governments mechanically default when expenditure shocks push the actual debt level beyond this limit. Unlike other studies, I assume perfect competition between investors and neglect moral hazard. One consequence is that self-fulfilling prophecies do not take place. Another consequence is an explicit premium function along with analytical properties, contributions to the literature.

Under certain conditions, I show that this premium is convex increasing in the actual debt level and convex decreasing in the debt capacity level, which depends on expected future fiscal surpluses. The result is consistent with the intuition. For instance, at the same debt level, the higher the debt capacity level, the lower the premium, and the more so the debt capacity level.

Numerical simulations show that fiscal consolidation and Eurobonds are both successful in improving public debt sustainability in highly indebted countries over the long run. However, private consumption and welfare effects differ. Because fiscal consolidations have a multiplier effect - higher fiscal surpluses also decrease the premium - they are welfare improving. Eurobonds on the other hand decrease welfare in highly-indebted countries: higher public debt draws from productive capital at home, creates an interest rate wedge between countries which attracts more investments from foreign countries, a trade re-balancing effect which builds foreign debt and depletes household wealth, welfare and consumption over the long run.

Further simulations show that Eurobonds can have a sizable cross-country insurance effect, when coupled with labor market reforms, if the premium across countries start from very different levels. Although inappropriate for quantitative statements, simulations finally indicate that Eurobonds have a positive short-run welfare effects if associated with tax reductions, but remain unable to compensate for the long-run losses under realistic valuation of the welfare of future generations.
Chapter 1

Flexicurity and Job Reallocation

Thomas Davoine and Christian Keuschnigg

April 2010, Updates (Thomas Davoine) June 2013

Abstract

This paper develops a general equilibrium model with safe and risky jobs where unemployment is concentrated in an innovative but volatile sector. Frictional unemployment arises in the process of job creation, firing and retraining for alternative employment. The paper derives an optimal welfare policy which combines the design of the tax schedule with three pillars of the ‘flexicurity’ model. The optimal policy is characterized by (i) a progressive wage tax schedule; (ii) a wage subsidy to re-employed workers; (iii) partial unemployment insurance benefits; (iv) job protection to contain firing; and (v) active labor market policy to facilitate labor reallocation, under certain conditions.

JEL-Classification: J64, J65, J68, J32, H30.

Keywords: Flexicurity, insurance, job protection, active labor market policy.

1We are grateful to the University of St.Gallen for financial support via the research program on ‘Risk and Wealth’ and for comments by seminars and workshops participants at the University of Konstanz, the University of St.Gallen, the 2010 CESifo area conference on Employment and Social Protection in Munich, the 2010 LMU Public Finance seminar in Munich, the 2011 Ruhr Graduate School conference in Dortmund, the 2011 L.A. Gérard-Varet conference in Marseille and the 2012 annual conference of the International Institute of Public Finance in Dresden.
CHAPTER 1. FLEXICURITY AND JOB REALLOCATION

1 Introduction

For decades, high unemployment has plagued welfare states especially in Europe. The causes of unemployment are manifold, and each one probably requires its own remedy. It is often argued that increasing globalization and the faster pace of technological progress lead to more volatile employment relationships, shorter job tenure and an increasing need for retraining of previously acquired skills (e.g. Brown, Merkl and Snower, 2009a, and Ljungqvist and Sargent, 1998). Part of unemployment thus results from an increasing speed of labor reallocation across different tasks with different skill requirements. Given the need to facilitate and speed up reallocation towards alternative employment, a successful policy might be the flexicurity model consisting of three pillars: insurance of the unemployed, active labor market policy (ALMP) to speed up transition into new jobs, and firing flexibility to close down unproductive jobs and replace them with new ones. Denmark’s success in reducing its unemployment rate from about 10% to 5% over the 1990’s is often attributed in good part to the flexicurity model.

The three pillars of flexicurity address separate channels of the labor market impact of welfare policies. Unemployment insurance (UI) is a central pillar of the welfare state and addresses a market failure due to missing private insurance markets for labor income risk. Developed countries spend up to 2% of GDP on UI (see OECD, 2009). Gruber (1997) estimated that the reduction in an unemployed workers’ consumption would be three times larger without UI (a 22.2% drop instead of 6.8%). Using a ‘sufficient statistics’ approach that avoids functional form assumptions, Chetty (2008) estimated that the current level of UI (replacement rate near 50%) is close to optimal in the US. There is consensus on the negative sign of the effect of UI on employment, but less so on its magnitude (see Holmlund, 1998). In their survey, Krueger and Meyer (2002) consider that a fair summary value for the effect of UI on employment is an elasticity of unemployment duration with respect to benefits of 0.5. There seems to be no or a positive impact of UI on wages or reservation wages.\(^2\)

The most immediate effect of employment protection (EP), independent of whether it comes as an administrative cost or as a firing tax, is to reduce job separation and, thereby, contribute to lower unemployment. On the negative side, EP makes firms reluctant to hire workers in the first place if they find it difficult and costly to separate again when prospects become unfavorable. For this reason, high levels of EP have long been blamed for high European unemployment rates. The flexicurity model advises a low level of EP, at least lower than in many generous European welfare states, since it is argued that the positive effect of more hiring on the unemployment rate outweighs the negative effect of a higher separation rate.

\(^2\)Ten studies on wage effects (5 reported by Addison and Blackburn, 2000, and the study itself, plus 3 reported by Petrongolo, 2008, and the study itself) arrive at different conclusions. They either find no statistically significant effect (in 6 cases) or statistically positive effects (higher UI benefits raise wages) which are often moderate or significant only for some subgroups (in the 4 other cases). Higher UI either raises wages or has no effect. The theoretical model below excludes wage effects.
This ambiguity is reflected in empirical research on how EP affects the labor market. Empirical studies consistently find that EP reduces flows into unemployment, but fail to report a reliable effect of EP on employment levels which points to a negative effect on hiring. The survey by Addison and Teixeira (2003) documents the heterogeneity of empirical findings since the original study by Lazear (1990). For instance, the summary table in Boeri and Jimeno (2005) shows only 5 significant coefficients out of 24. Heckman and Pages (2000) find significant positive effects of EP on unemployment. They estimate that the average 3 months of firing costs in Latin America amounts to a loss of 5.5% points of employment (for an average of 7.4% of unemployment). On the other end, OECD (1999) reports no significant effects of EP. Based on their empirical estimates, Belot, Boone and van Ours (2007) calculate an optimal level of employment protection which is positive and moderate: for instance, the growth maximizing level of protection for open-ended contracts is around 0.37 (on a scale from 0 to 1), corresponding to 1999 levels of protection in countries like Italy, Switzerland or the UK.

Active labor market policy (ALMP) mostly aims to support job search effort of unemployed workers, to develop their skills and make them more attractive to potential employers, and to retrain for different jobs with alternative skill requirements. Total spending on labor market policies has grown to significant levels over the years. According to OECD (2009), member countries spent on average 1.5 % of GDP in 2006 with some countries spending up to 3.4%. The fraction devoted to some form of active measures has grown from 35% in 1985 to 42% in 2007 (Martin and Grubb, 2001; and OECD, 2009). Several empirical studies show that monitoring and sanctions increase search efforts by unemployed workers more than the reduction in unemployment benefits (reported by Nunziata, 2008). Early empirical studies find negative or no significant effects of ALMP training programmes in the short-run (e.g. see the surveys Heckman, Lalonde and Smith, 1999; Martin and Grubb, 2001; Kluve, 2006), mostly due to a lock-in effect, but recent studies have data to focus on long-run effects. For instance, Lechner, Miquel and Wunsch (2011) find that the re-employment probability increases by 20 to 40% and monthly earnings are higher by 0 to 550 Euro, depending on the training type.

There has been extensive theoretical research on different causes of unemployment and ways to reduce it. Economists have often studied the effects of different policies in isolation or in pairs. Using a combination of instruments makes theoretical models more complicated. However, it is important to analyze all three pillars simultaneously to capture the full potential of the flexicurity model as well as the interactions and complementarities between labor market policies and the wage tax schedule. Andersen and Svarer (2007) argue that low EP alone does not explain the decline of unemployment in Denmark. Low EP and generous UI were already in place well before the rise in unemployment, following the mid-1970’s oil shock. Only when Denmark implemented activation measures for the unemployed, did unemployment start to come down.

A simple cross-country comparison further illustrates the potential of flexicurity. In international comparison, both Denmark and (West) Germany had large and increasing
ALMP programmes in the 1990’s. However, Denmark had more generous UI and lower EP. The resulting flexicurity combination was associated with a decline in unemployment over the 1990’s, while unemployment raised in (West) Germany.3

Most of the previous theoretical work includes some, but not all of the three policy instruments. For instance, some papers consider EP and UI together (such as Pissarides 2001; Blanchard and Tirole 2008; Cahuc and Zylberberg, 2008), but do not include ALMP. In this vein, Blanchard and Tirole (2008) show that it may be preferable to finance UI with firing taxes rather than wage taxes or contributions. However, they neither include job creation nor ALMP. Some of the macroeconomics literature assumes risk-neutral workers to isolate the incentive effects of UI but misses gains from insurance which are, after all, the prime motivation for providing social insurance in the first place. Most theoretical literature on ALMP also considers UI (reviewed in Fredriksson and Holmlund, 2006). Even though some of these ALMP and UI papers have other policy instruments (e.g. welfare in Pavoni and Violante, 2007), none explicitly includes EP.

Three recent papers focus on the flexicurity model but do not explore all possibilities afforded by the three policy instruments. Andersen and Svarer (2008) consider UI and ALMP but assume policy makers commit to no EP. With a simulation, they show that using workfare (ALMP) may be one way to improve labor market performance without reducing UI benefits. Brown, Merkl and Snower (2009b) limit EP to firing costs and do not use it as a firing tax which could be used to finance UI as suggested in Blanchard and Tirole (2008) and which is, in fact, partly implemented in the U.S., for example. In Brown, Merkl and Snower (2009b), one can also note that numerical results are sensitive to one parameter (the probability that a firm hires temporary workers from a secondary labor market when its own workers go on strike due to a too low wage offer) which has never been estimated and, arguably, is unusual in the wage bargaining process. With this caveat in mind, their simulation shows that unemployment in Germany could be reduced by 50% if it adopted the same UI, ALMP and EP policies as Denmark. Algan and Cahuc (2009) do not explicitly analyze ALMP. Extending the Blanchard and Tirole (2008) framework to the case of moral hazard, they show that only countries with high levels of civic attitude would benefit from flexicurity, as in Denmark.

Theoretical studies of EP have reached contradictory conclusions, too: early studies with EP alone find opposing effects in simulation exercises. Bertola (1990) as well as Blanchard and Portugal (2001) show that higher EP can boost employment, under certain circumstances. Hopenhayn and Rogerson (1993) find the opposite. More recent studies with UI as a second policy instrument also arrive at different conclusions. Blanchard and Tirole (2008) conclude that firing taxes, not wage taxes should be used to finance UI.

---

3Specifically, Denmark spent 1.66% of GDP on ALMP in 1998, had UI with an average replacement rate of 66% in 1999 and an EP index of 0.70 out of a maximum 2 in the same year. Its average unemployment rate declined from 8.1% in 1988-1995 to 5.3% in 1996-1999. Corresponding numbers for (West) Germany were 1.26% of GDP for ALMP, 37% for UI, 1.30 for EP, 5.6% and 7.1% for unemployment rates (numbers taken from Nickell & all, 2005). This paper being theoretical, causal empirical analysis is left for future research.
1. INTRODUCTION

and EP should be as high as necessary to internalize firing externalities. Andersen and Svarer (2008) assume flexible hiring and firing by keeping EP at zero. Unlike Gruber (1997), but for a value closer to Chetty (2008), Andersen and Svarer (2008) show that unemployment can still be reduced with values of the replacement rate that would be considered high in the theoretical literature (around 0.5).

Models with few instruments have the virtue of simplicity and unambiguous theoretical predictions but cannot capture the full meaning of a flexicurity policy. The aim of this paper is to investigate the joint effects of the three pillars of the flexicurity model and rationalize it as an optimal policy outcome together with the design of the tax benefit scheme in labor income taxation. A second benefit of our model is that we can look at factor supply and efficient utilization at the same time while other models do it separately. For instance, the Baily (1978) family of papers (including Gruber, 1997, and Chetty, 2008) look at the impact of UI on search behavior which is an aspect of labor supply while Acemoglu and Shimer (1999) investigate the impact of UI on the job match quality and thereby explain labor productivity by the efficiency in the allocation of labor. But each stream does it separately. A third novel feature of our model is the distinction of more and less volatile sectors with differing incidence of unemployment, an idea borrowed from Cunat and Melitz (2007) which makes the aggregate unemployment rate a function of the economy’s sectoral composition. We specifically introduce the notion of retraining and job search to show how an optimal welfare policy should facilitate ongoing structural change and support the reallocation of labor. Similar to the probabilistic modeling of education investments in Konrad (2001), we assume that retraining of sector specific skills and search for re-employment in the second sector is risky. A larger effort spent on retraining and job search raises the likelihood to find employment in the alternative sector. If unsuccessful, the worker remains unemployed and collects benefits.

Our analysis of job reallocation is also related to the endogenous growth and labor market literature. Aghion and Howitt (1994) embed a search and matching model in a Schumpeterian growth model and show that creative destruction can lead to higher unemployment, when the rate of innovation increases. Instead of search and matching, Grossmann (2000) adds labor market segmentation and an efficiency wage setting to the Schumpeterian growth model and arrives at comparable results. We abstract from endogenous growth and use a static model but add policy analysis.

This paper presents a model with three instruments of labor market policy, UI, ALMP and EP, together with a possibly progressive wage tax schedule including tax credits and wage subsidies to support retraining. In line with empirical evidence, we assume that EP directly affects hiring and firing of firms while UI and ALMP affect retraining. In line with empirical evidence, we assume that EP directly affects hiring and firing of firms while UI and ALMP affect retraining.

---

4At the firm level, Comin and Philippon (2005) and references in this paper show that firm volatility is a ‘good predictor of both unemployment risk and wage inequality’. At the industry level, Davis, Faberman, Haltiwanger, Jarmin and Miranda (2008) find results that ‘supports the view that industry differences in the intensity of idiosyncratic shocks are a major reason for industry differences in the incidence of unemployment’, where industry volatility (idiosyncratic shock) is measured by variation of firm size (number of employees).
job search efforts of dismissed workers. Search effort as well as job creation and job destruction are endogenously defined, the last two being driven by productivity shocks. To address the importance of endogenous job reallocation in a volatile environment, the model uses two sectors with *retraining* of skills after separation and *reemployment* of workers in another sector. The first sector may be thought of as an innovative industry with volatile employment relationships. In the event of separation from a sector 1 job, workers can retrain and search for a job in a second sector offering safe employment relationships. In the spirit of Ljungqvist and Sargent (1998), we assume that job separation leads to a loss of (sector specific) skills so that the next best job pays a lower wage, which is empirically supported.\(^5\) Once a job is obtained in this sector, the worker is never fired again. Upon changing sectors, adjustment of skills are necessary; a worker who entered the first sector but got fired, needs to spend effort on retraining and search to find a job in the second sector. The amount of search effort and, in turn, successful job reallocation, depends on the levels of ALMP and UI. The last assumption refers to the welfare increasing role of ALMP. These services might partly have the characteristics of a public good (e.g. increasing market transparency and providing information about job opportunities to reduce private search costs), or they might be a publicly provided private good (training services, consulting to improve job applications etc.) for which markets do not exist for reasons outside the model.\(^6\)

To sum up, this paper analyzes the design of the wage tax schedule jointly with a ‘flexicurity’ policy, consisting of flexibility in job separation, UI and ALMP. Apart from redistributive goals, the policy instruments need to address three market failures: missing markets for UI, firing externalities, and absence of private supply of job search assistance. Our analysis yields five results on an optimal tax and flexicurity program. First, tax progressivity is motivated by redistribution from workers in highly paid volatile jobs to those with lower wages in more stable employment relationships. To maintain incentives optimal redistribution is partial. Second, the tax schedule is complemented by tax credits or subsidies to encourage retraining and search for alternative employment. Third, public UI is necessary due to missing private markets. Again, to maintain incentives optimal insurance is partial. Fourth, negative firing externalities create a need for an optimal degree of job protection. And fifth, active labor market policy is an essential ingredient of a large welfare state if its costs are low, it is sufficiently productive in stimulating reemployment and raising individual welfare by reducing private search costs. On the other hand, effects on unemployment are most of the time ambiguous. As expected

\(^5\) Jacobson, LaLonde and Sullivan (1993) find that after a mass lay-off in a distressed economic environment, workers settle for wages that are 25% lower (6 years after the lay-off, in the new job). In a less turbulent environment, Couch and Placzek (2010) find that wages are 12 to 15% lower.

\(^6\) An unproductive, sanctions based ALMP cannot be part of an optimal policy in our framework. Although effective in raising job search, these policies reduce private utility and are, in addition, costly to the government. Only a productive, welfare increasing form of ALMP can be rationalized. However, when workers are heterogeneous and job search behavior is private information, sanctions might possibly be a way to separate different worker types similar workfare requirements as in Besley and Coate (1995) and Kreiner and Traenes (2005) and might thus become part of a welfare optimal program.
UI increases unemployment. Tax progressivity however helps curb unemployment, as reduced taxes in the stable employment relationships motivates job seekers.

The next section sets up the model, Section 3 derives optimal policies, Section 4 considers piecemeal reform and Section 5 illustrates numerically how optimal policies adjust to changing structural parameters. Section 6 concludes.

2 The Model

2.1 Sectoral Production

The economy is populated by a mass $N$ of risk-averse workers and consists of two sectors producing the same numeraire good. The first sector is made of new firms and the second sector of existing firms. New firms introduce innovative production processes and includes start-up firms, self-employed workers and all spin-offs from existing firms, so the first sector can be relatively large. The innovation outcome is uncertain: if successful, the new firm is more productive than existing firms; if unsuccessful, the firm is less productive and needs to close down. The first sector is thus a volatile industry, more productive on average but with larger job turnover and unemployment risk. Sector 2 is less productive, pays lower wages but offers safe jobs.

Set up costs and set up time are needed to start new firms (see Fonseca et al., 2001). Set up costs are borne by firm owners and set up time by workers. Before producing the good, workers need to set up the production process. For instance, the oven needs to be purchased by the owner of a new bakery and installed by the worker. Then, just as existing bakeries, the oven is operated by the worker. Workers are willing to invest extra time in new firms because the production process has a chance to be more efficient and, in this case, they can collect higher wages.

A part $N$ of workers is ready to invest extra time and seeks employment in the volatile industry. The remaining part $1 - N$ does not invest and accepts a lower paying, but safe job in sector 2. Initially, the sectoral allocation of labor results from occupational choice with a discrete time investment. Before entering the labor market, workers train to acquire the skills needed for their target occupation, whether in the first or the second sector. After a productivity shock, a share of workers in the volatile industry is fired because the job turned out unproductive. In this case, part of the workers target the same occupation in sector 2 and another part a different occupation. Some unsuccessful self-employed lawyers decide to become bankers in existing firms in sector 2. In average, sector 1 fired workers need to acquire different skills. Fired workers can retrain and search for a sector 2 job. All this is anticipated when making the initial choice of sector.

Figure 1 illustrates how retraining leads to reallocation of labor. When the outcome of the productivity shock is unfavorable, employment in sector 1 is terminated, leading to job separation with probability $s$ and continuation with probability $1 - s$. When fired,
the worker can retrain and search for a sector 2 job. When job search is not successful, she remains unemployed. When entering the volatile industry, a worker may thus end up in three states. They ultimately keep their job in sector 1 with probability $1 - s$, are reallocated to jobs in sector 2 with probability $se$, and end up unemployed with probability $s(1 - e)$. Given independent risks, the ex ante probabilities correspond to ex post fractions. Initial and final labor allocation must satisfy the resource constraint,

$$L_1 = (1 - s)N, \quad L_2 = (1 - N) + esN, \quad \delta \equiv (1 - e)sN = 1 - L_1 - L_2. \quad (1)$$

The unemployment rate $\delta$ reflects job creation $N$, firing $s$ in the volatile industry and unsuccessful job search $1 - e$ for new employment. Since the unemployment risk is high in the volatile sector 1 and low (zero) in sector 2, the average unemployment rate necessarily reflects the sectoral composition of the economy.

Successful new firms are more productive and share a part of their profits with workers, so workers who keep a job in sector 1 earn a higher wage, $w_1 > w_2$. Separation and reallocation leads to a wage loss, so retrained and re-employed workers have a lower wage, $w_r < w_1$. We further assume that sector 2 technology is Ricardian with fixed productivities $w_r > w_2$: workers separated from a sector 1 job are more productive than those who started a job in sector 2 right from the beginning, because they accumulated some additional experience. After reallocation, employment in sector 2 consists of $1 - N$ unskilled and $esN$ retrained workers, earning fixed wages $w_2$ and $w_r$, respectively. If a sector 1 worker becomes unemployed, she generates low income $h$ from home production. To sum up, we assume $w_1 > w_r > w_2 > h$, where the lower index $r$ indicates re-training.

We consider social protection in the context of a possibly progressive wage tax schedule and allow for different proportional tax rates $t_1$, $t_r$ and $t_2$ in each earnings class with wages $w_1 > w_r > w_2$. While $t_1 > t_2$ naturally describes a progressive tax schedule since sector 1 workers on average earn more. Whenever $t_r < t_1$, we associate the difference $t_1 - t_r$ with a wage subsidy\(^7\). A priori, these tax rates are unrestricted. In addition, the government sets unemployment benefits $b$, may impose a firing tax $t_s$ to reduce job separation and spend on ALMP $m$ to support retraining and assist job search\(^8\). Taking policy instruments as given, workers decide whether to invest set-up time in sector 1 and, in the event of separation, choose a level of retraining and job search effort. Firms decide whether to employ a worker and, after the productivity shock materializes, whether to close down or continue the employment relationship.

Sector 1 production is organized by risk-neutral firms, each hiring one worker. We assume owners have a portfolio of firms. With perfect competition, firm entry and job

\(^7\)To investigate the effect of progressive tax schedule and wage subsidies, differences in tax rates $t_1$, $t_r$ and $t_2$ are needed. It thus makes sense to make a difference between wages $w_1$, $w_r$ and $w_2$. In particular, we refrain from equating $w_r$ and $w_1$ and consider that prior experience in sector 1 provides some productivity benefit over direct entry into sector 2, so that $w_r > w_1$.

\(^8\)Severance payments are more common than firing taxes as a mean of employment protection. Among the OECD countries, only the US is making use of firing taxes (via the so-called experience rating system for financing unemployment insurance; see OECD, 2004, for more). Even though not widespread, firing taxes remain a realistic policy instrument.
creation continue until profits are zero. After hiring, firms are subject to a productivity shock \( x \in \left[0, \infty\right) \), leading to output \( Ax \) of the job. Once the firm is set up and thus productivity is known, the firm decides whether to continue with earnings \( Ax - w_1 \) or close down, fire the worker and accept a loss \( t_s \) equal to the firing tax. In this case, The firm continues if \( Ax - w_1 \geq -t_s \). The cut-off productivity is

\[
x_1 = \frac{(w_1 - t_s)}{A}.
\]  

(2)

When the productivity shock yields a better result \( x \geq x_1 \), the firm continues the employment relationship, in the other case, the job is terminated. Given a density \( g(x) \) and cumulative distribution \( G(x) \), the separation rate is

\[
s(x_1) = \int_0^{x_1} g(x) \, dx, \quad 1 - s(x_1) \equiv \int_{x_1}^{\infty} g(x) \, dx.
\]  

(3)

A higher cut-off value \( x_1 \) raises the firing rate \( s \) and reduces the continuation probability \( 1 - s \). Note that \( 1/s \) is interpreted as the length of job tenure in the volatile industry. High volatility means a high firing rate and short job duration, i.e. high job turnover.

Entry and job creation give rise to the fixed cost or start-up investment \( f \). Since owners start with several firms, they can pay sunk investment costs and firing taxes for these firms which need to close down with the profit they make with successful firms. Anticipating the firing decision, firms create jobs if the net present value is non-negative. Defining average productivity after entry by \( x^a \equiv \int_{x_1}^{\infty} x dG(x) \) \( / (1 - s) \), expected profits

Figure 1: Reallocation of Labor
CHAPTER 1. FLEXICURITY AND JOB REALLOCATION

Employment relations also break up when a worker prefers to leave. To prevent this, the firm must pay a high enough wage. Utility of staying in the firm is $u((1-t_1)w_1)$. Job creation under perfect competition pushes up the wage until profits $\pi$ are zero. When firing is optimally chosen as in (2), the derivative of the profit function with respect to cut-off productivity is zero, $d\pi = -(Ax_1 - w_1 + ts)g(x_1)dx_1 = 0$. By the envelope theorem, $d\pi/dw_1 = -(1-s)$ and $d\pi/dts = -s$. Hence, the zero profit condition pins down the competitive wage as a function of the firing tax and other fundamental parameters. Solving $d\pi = -(1-s)dw_1 - sdt_1 = 0$ yields

$$
\frac{dw_1}{dt_1} = -\frac{s}{1-s}, \quad \frac{ds}{dt_1} = -\sigma s, \quad \sigma \equiv \frac{g(x_1)}{(1-s)sA}.
$$

Hence, a firing tax puts a cost on firms and forces them, in zero profit equilibrium, to cut the wage. For the same reason, the tax also reduces the separation rate, $s$. The semi-elasticity $\sigma$ measures the impact of firing taxes on separation rates.

Sector 2 uses a linear Ricardian technology. Competitive producers pay wages equal to exogenous labor productivities $w_2$ and $w_r$ of a specialized sector 2 worker and a retrained sector 1 worker, and earn zero profits. These wages are fixed constants while $w_1$ is endogenous. Since each firm hires exactly one worker, the number of new firms is determined by the number $N$ of workers entering sector 1. Since only $1-s$ new jobs survive, and $s$ close down, the number of productive jobs (or mature firms) is $(1-s)N$. Given average productivity $x^a$, total sector 1 output $X_1$ net of entry costs amounts to

$$
X_1 = [(1-s)Ax^a - f]N, \quad X_2 = w_2(1-N) + w_reN.
$$

Sector 2 output $X_2$ is produced by two types of workers, initial sector 2 and retrained sector 1 employees, each generating an output per capita equal to $w_2$ and $w_r$, respectively.

2.2 Labor Market Behavior

There are four realizations of income, $y \in \{w_1, w_r, h, w_2\}$. Net of tax income depends on the tax benefit schedule summarized by $\{t_1, t_2, tr, b\}$. As the model is static, net income is entirely consumed. Sector 1 and sector 2 goods being identical, household preferences are directly defined by net income, taking risk aversion and uncertainty into account. A sector 2 worker earns a safe wage, thus giving utility $V_2 = u((1-t_2)w_2)$. A sector 1 worker either keeps her initial job or is fired. When fired, she may retrain and get another
2. THE MODEL

job with probability $e$, or end up unemployed with probability $1 - e$ (see Konrad, 2001, for a probabilistic model of human capital investment). Taking account of a utility loss ( stigma ) $\chi$ of remaining unemployed, expected utility of entering sector 1 is

$$V_1 = (1 - s) \cdot u((1 - t_1) w_1) + s \cdot u^e,$$

$$u^e = \max_e e \cdot u((1 - t_r) w_r) + (1 - e) \cdot [u(h + b) - \chi] - \phi(m) \zeta(e).$$

Parameter $m$ represents the amount of active labor market policy ( ALMP ) provided, designed to facilitate retraining by reducing individual effort cost. The function $\phi$ represents the effectiveness of the ALMP policy. We assume $\phi(0) = 1$, $\phi' < 0 < \phi''$ and $\lim_{m \to \infty} \phi(m) = \phi_0 > 0$. When the policy is scaled up, it becomes less and less effective. For given $m$, effort costs are convex increasing, $\zeta' > 0$ and $\zeta'' > 0$, and are assumed strictly positive in the relevant range. The level search effort costs $\phi \zeta$ may also be interpreted as ‘stigma of job loss’ because it reduces utility value outside of sector 1 employment while $\chi$ reflects a ‘stigma of unemployment’. These two concepts are not the same since a fraction of separated workers successfully retrain but may not enjoy the new job to the same extent. Finally, the negative cross-derivative $\phi' \zeta'$ means that ALMP reduces marginal effort cost and stimulates job search.

The utility cost $\phi \zeta$ might also be interpreted as skill specificity since these individuals have previously invested for their initial occupation in sector 1 and must now retrain to obtain another job in sector 2. Another aspect of skill specificity is the assumption of $w_1 > w_r$. An initial sector 1 worker is not as productive in a sector 2 job after retraining and, as a result, must accept an earnings loss (see Ljungqvist and Sargent, 1998, on the loss of skills due to unemployment). The effort spent on retraining and job search for employment elsewhere is a matter of incentives. After job separation, individuals choose effort according to

$$u((1 - t_r) w_r) - u(h + b) + \chi = \phi(m) \zeta'(e).$$

Anticipating subsequent events, workers must decide in the beginning whether to undertake sector 1 set-up investment or go to sector 2 right away. Suppose agents are arranged by the innate ability $n \in [0, 1]$ for setting up firms. Ability reflects for instance entrepreneurial skills which can not be learned or appetite for effort. Depending on ability, set-up efforts are higher or lower. The effort cost $i(n)$ function represents the utility cost for the worker in setting up sector 1 firms and differs by ability according to $i'(n) > 0$, $i(0) = 0$ and $i(n) \to \infty$ for $n \to 1$, assuming continuity in the function $i(n)$. Low $n$ indicates low effort cost and high ability. Suppose ability is uniformly distributed so that the pivotal value $N \in [0, 1]$ is also the fraction of individuals with ability $n < N$. Given that expected utility of entering sector 1 is higher than that in sector 2, $V_1 > V_2$, highly able individuals with low cost expect $V_1 - i(n) > V_2$ and, thus, invest in new

---

9Blanchard and Tirole (2008) and Algan and Cahuc (2009) similarly assume a fixed utility cost of unemployment. Since they abstract from reemployment and allow for only one state after separation, they do not differentiate the utility loss from separation.
firms creation in sector 1. In the other case, type \( n \) opts for sector 2 and does not invest. The pivotal agent, identified by the occupational choice condition, determines the initial labor allocation across sectors,

\[
V_1 - i(N) = V_2, \quad V_2 = u((1 - t_2)w_2).
\] (9)

Initially, \( N \) workers opt for sector 1, and \( 1 - N \) go to sector 2. After firing, this allocation is partly revised by retraining, leading to a final allocation \( L_1 \) and \( L_2 \) as in (1). If there were no firing \( (s = 0) \), there would be no reallocation, \( L_1 = N \), and no unemployment. Unemployment results from productivity shocks, frictions in retraining and reallocation which only arise in the volatile industry.

### 2.3 Equilibrium

Government spends \( b\delta \) on unemployment benefits and \( mksN \) on ALMP, where \( m \) is the amount of ALMP provided, \( k \) is the unitary cost of ALMP and thus \( mk \) is spending per capita of fired persons in need of a new job. Fiscal budget balance requires

\[
T = t_2w_2(1 - N) + t_1w_1(1 - s)N + t_rw_1esN + tsN - b(1 - e)sN - mksN - C = 0, \quad (10)
\]

where \( C \) is an exogenous and constant level of other public spending.

Aggregate disposable income stems from earnings of employed and retrained sector 1 workers (first two terms below), benefits collected by unemployed persons, and earnings of specialized sector 2 workers. Since \( h \) is income from non-market activity, it does not show up in disposable income,

\[
Y = (1 - t_1)w_1(1 - s)N + (1 - t_r)w_1esN + b(1 - e)sN + (1 - t_2)w_2(1 - N). \quad (11)
\]

Using the fiscal budget to replace benefits, profits \( \pi = (1 - s)(Ax^a - w_1) - st_e - f \) to eliminate \( w_1 \) and substituting the definitions \( X_j \) yields \( (Y + C + mksN - X_1 - X_2) + \pi N + T = 0 \), where the bracket is excess demand. Total demand for market goods includes not only private and public consumption \( Y + C \) but also the resource use of ALMP spending. Solving for \( \pi = T = 0 \) clears the product market by Walras’ Law.

The solution of the untaxed, free market equilibrium is as follows. Parameters \( w_r \) and \( h \) determine expected utility if fired, \( u^e \). Suppose that the lowest possible sector 1 wage, \( u(w_1^*) = u^e \), implies positive expected profits \( \pi(w_1^*) > 0 \), reflecting high factor productivity \( A \) and average productivity \( x^a \). This wage also induces a labor allocation \( N^* \), satisfying \( V(w_1^*) - i(N^*) = u(w_2) \). Profits attract new firms which bid up the wage rate \( w_1 > w_1^* \). Expected utility \( V_1 \) rises, attracting more workers such that \( N > N^* \) satisfying \( V(w_1) - i(N) = u(w_2) \). This process continues until profits are zero, \( \pi(w_1) = 0 \). At this equilibrium wage, the participation constraint is not binding, \( u(w_1) > u^e \). We ‘calibrate’ the model such that \( w_1 > w_r > w_2 > h \).
3. Optimal Flexicurity

In the present model, policy should address three market distortions, arising from firing externalities, missing private insurance markets, and frictions in retraining and job search. ALMP could be interpreted as a non-rival public good providing market transparency and public information about employment opportunities that are useful to guide individual training and job search effort and to facilitate transition into new jobs. Or it could be a publicly provided private good such as advice on how to apply for jobs, or subsidized training if new skills are required in alternative occupations. We first analyze how firms and households react to policy changes and then characterize optimal policy.

To obtain compact formulas, we define notation

\[ u_1 \equiv u((1 - t_1)w_1), \quad u_r \equiv u((1 - t_r)w_r), \quad u_h \equiv u(h + b) \] as well as \[ u_2 \equiv u((1 - t_2)w_2). \]

The index refers to the final state of the worker, employed in sector 1, re-trained and employed in sector 2, engaged in home production while unemployed or employed in sector 2.

3.1 Behavioral Effects and Fiscal Impact

Wages and separation rates in sector 1 depend exclusively on the firing tax as in (5), but are independent of the tax benefit schedule for workers. Taking the differential of (7-8) reveals the policy impact on expected utility \( u_e \) and job search after separation,

\[
du_e = -u_r' \cdot w_r \cdot ed + u_h' \cdot (1 - e) \cdot db - \phi' \zeta' \cdot dm, \tag{12}
\]

\[ de = -\epsilon_r \cdot w_r \cdot edr - \epsilon_b \cdot (1 - e) \cdot db + \epsilon_m \cdot dm, \]

where behavioral elasticities are all defined positive,

\[ \epsilon_r \equiv \frac{u_r'}{e \phi' \zeta''} > 0, \quad \epsilon_b \equiv \frac{u_h'}{(1 - e) \phi' \zeta''} > 0, \quad \epsilon_m \equiv -\frac{\phi' \zeta'}{\phi' \zeta''} > 0. \]

In reducing private effort costs, ALMP spending boosts retraining and job search. Taxes and benefits discourage job search and retraining. A higher labor tax \( dt_r > 0 \) reduces re-employment \( e \) because it makes work less attractive. As reflected in the elasticity \( \epsilon_r \), the impact is larger when the marginal utility of re-employment \( u_r' \) is large and the job search assistance impact \( \phi \) is low: in the latter case, worst consequences of firing reduce entry in sector 1 and thus the flow out of sector 1 into unemployment. Higher unemployment benefits \( db > 0 \) also reduce re-employment, as it reduces search incentives. The reduction in search incentives is magnified by the utility costs \( u_h' \) and, as before, low impact of job search assistance \( \phi \), both reflected in the elasticity \( \epsilon_b \). Higher job assistance \( dm > 0 \) on the other hand increases re-employment flows \( e \). The strength of the effect, the elasticity \( \epsilon_m \), depends on the marginal benefit of job search assistance and search effectiveness, \( \phi' \zeta' \).

The tax distortions are further measured by the participation tax \( \tau^E \equiv t_r w_r + b \) which will be introduced below and consists of the sum of the wage tax plus the benefits lost.
when an individual switches from unemployment into a job. Expected utility rises with more generous benefit levels and a more intensive ALMP but falls with a higher tax burden on employment.

At the beginning, when seeking employment in sector 1, agents anticipate the separation risk and expect utility \( V_1 = (1-s)u_1 + su^e \). Using \( \nabla \equiv (u_1 - u^e) / u'_1 \),

\[
dV_1 = u'_h (1-e) s \text{d}b - u'_r w_r s \text{d}t_r - \phi \xi s \text{d}m
- u'_1 w_1 (1-s) dt_1 - (1-t_1 - \nabla \sigma) u'_1 s \text{d}t_s. \tag{13}
\]

After starting employment in the volatile industry, individuals expect to be fired and remain unemployed with probability \((1-e)s\). Expected utility rises by the marginal welfare gain \( u'_1 db \) from more generous benefits, scaled by the probability of unemployment. Other taxes and benefits are interpreted similarly. More intensive ALMP becomes useful in the event of job separation which is expected with probability \(s\). Importantly, a higher firing tax \( dt_s \) affects workers via two offsetting channels and leaves an a priori ambiguous net effect. On the one hand, the tax reduces firing by \( ds = -\sigma s \text{d}t_s \) which boosts expected utility by \( \nabla u'_1 (-ds) = (u_1 - u^e) (-ds) \). Less firing allows workers to enjoy more often high utility \( u_1 \) from sector 1 employment instead of low utility \( u^e \) as expected after separation. On the other hand, the tax reduces the gross wage by \( dw_1 = -\frac{s}{1-s} dt_s \). Since this occurs with probability \( 1-s \) ex ante, expected welfare declines by \( (1-s) u'_1 (1-t_1) dw_1 \).

The willingness to pursue employment in the volatile industry and invest in firms set-up depends on expected utility \( V_1 \) relative to welfare \( V_2 = u ((1-t_2) w_2) \) from a safe sector 2 job. Taking the differential of the occupational choice condition in (9) yields the entry response,

\[
dN = -\eta_1 w_1 (1-s) N dt_1 - \eta_r w_r s N \text{d}t_r + \eta_h (1-e) s N \text{d}b
- \eta_s s N \text{d}t_s + \eta_m s N \text{d}m + \eta_2 w_2 (1-N) dt_2, \tag{14}
\]

where entry elasticities are defined positive,

\[
\eta_1 \equiv \frac{u'_1}{N t'}, \quad \eta_2 \equiv \frac{u'_2}{(1-N) i'}, \quad \eta_m \equiv -\frac{\phi \xi}{N t'},
\]

\[
\eta_r \equiv \frac{u'_r}{N t'}, \quad \eta_h \equiv \frac{u'_h}{N t'}, \quad \eta_s \equiv (1-t_1 - \nabla \sigma) \eta_1.
\]

A higher tax \( dt_2 \) on sector 2 earnings pushes workers into sector 1 employment. The elasticity \( \eta_2 \) measures the strength of the effect: the lower the increase in sector 1 firm creation investment \( i' \) and the higher sector 2 utility costs \( u'_2 \), the higher the impact on sector 1 enrollment. Conversely, all policies raising the present value of net taxes on sector 1 earnings discourages employment in the volatile sector. Clearly, higher unemployment benefits \( db \) reduce the present value of net taxes and, thus, make volatile

\footnote{The participation tax is implicit in (8). To see this, substitute the Taylor expansion \( u_r = u_h + u'_h (1-t_r) w_r - b - h \) and note the definition of \( \tau^E \), yielding \( w_r - \tau^E - h \) \( u'_h + \chi = \phi \xi' (e) \).}
employment more acceptable. The higher the utility gains of unemployment $u'_h$, the stronger the effect, as reflected in the elasticity $\eta_h$. More intensive ALMP $dm$ improves labor market prospects in the event of job separation and, thus, similarly encourages sector 1 employment. The higher the marginal increase in search assistance impacts labor market prospects in the event of job separation and, thus, similarly encourages sector 1 employment. To make sense of elasticity $\eta_s$, simply note that $\eta_s$ is expressed in terms of the elasticity $\eta_1$ in the same fashion that changes in firing taxes $dt_s$ are expressed in terms of changes of labor taxes $dt_1$ in expected utility $dV_1$, see (13).

Employment effects on different margins determine the tax yield and the net result on the fiscal constraint in (10). In the present model, all labor market effects are discrete in the sense of people switching from one state to another, either via discrete firm creation investment $dN$, job separation $ds$ or retraining with successful job search $de$. For each of these margins, we define effective tax rates $\tau^N$, $\tau^S$ and $\tau^E$, which capture the impact of behavioral changes on net tax revenue,

$$dT = \left(1 - N\right) \cdot w_2 dt_2 + \left(1 - s\right) N \cdot w_1 dt_1 + esN \cdot w_r dt_r + \left(1 - t_1\right) sN \cdot dt_s - \left(1 - e\right) sN \cdot db - ksN \cdot dm + \tau^N \cdot dN + \tau^S \cdot N ds + \tau^E \cdot sN de,$$

where effective tax rates on the extensive margins of employment are defined as

$$\tau^E \equiv t_r w_r + b, \quad \tau^S \equiv t_s + \left[et_r w_r - (1 - e) b\right] - km - t_1 w_1, \quad \tau^N \equiv t_1 w_1 - t_2 w_2 + s \tau^S.$$

$\tau^E$ represents the effective tax rate on labor market participation, $\tau^S$ the effective tax rate on firing and $\tau^N$ the effective tax rate on sector 1 entry. Using these rates, we can write the government budget constraint as $T = t_2 w_2 + \tau^N N = 0$. If one more person switches from sector 2 employment into sector 1, net tax revenue rises by $\tau^N$. The net impact consists of the differential tax liability of sector 1 over sector 2 employees $t_1 w_1 - t_2 w_2$ plus the additional net tax revenue $\tau^S$ that is collected when a sector 1 worker gets fired, an event which occurs with probability $s$. The ‘effective’ firing tax $\tau^S$ captures the fiscal consequences of job separation. It consists of the firing tax $t_s$ paid by firms, plus the average net tax liability of a worker after separation, equal to $et_r w_r - \left(1 - e\right) b$, minus spending $km$ on ALMP per capita of a fired worker, minus the foregone tax $t_1 w_1$ when this person is no longer employed in sector 1. Writing the net tax liability after separation as $et_r w_r - \left(1 - e\right) b = e \tau^E - b$ reveals the fiscal gain $\tau^E$ of putting one more person back to work, consisting of the wage tax $t_r w_r$ of this reemployed person plus the savings in UI spending when the same person no longer collects benefits $b$.

Using these effective tax rates and substituting the behavioral changes given above yields a change in net fiscal revenue equal to

$$dT = \left(1 + \tau^N \eta_2\right) w_2 \left(1 - N\right) dt_2 + \left(1 - \tau^N \eta_1\right) w_1 \left(1 - s\right) N dt_1 + \left(1 - \tau^N \eta_r - \tau^E \epsilon_r\right) w_r esN dt_r - \left(1 - \tau^N \eta_h + \tau^E \epsilon_h\right) \left(1 - e\right) sN db + \left(1 - t_1 - \tau^S \sigma - \tau^N \eta_s\right) sN dt_s - \left(k - \tau^N \eta_m - \tau^E \epsilon_m\right) sN dm. \quad (16)$$

For example, spending on more intensive ALMP rises by $k \cdot sN dm$ which obviously is a loss of net tax revenue. Since the policy attracts $\eta_m sN dm$ more workers to sector 1 as in
CHAPTER 1. FLEXICURITY AND JOB REALLOCATION

(14), and each one adds \( \tau N \) in expected value to the government budget, net tax revenue rises by \( \tau N \eta msNd \). The budget further improves since ALMP brings a larger portion of all fired persons back to work and raises the number of reemployed sector 1 workers by \( \varepsilon msNd \). Each of these persons who were previously unemployed and now get a job, pays tax and stops claiming benefits, adding \( \tau E \) to the budget. Adding up all these consequences, the net fiscal cost of more intensive ALMP is only \( (k - \tau N \eta m - \tau E \varepsilon m) sNd \) instead of \( ksNd \). The difference reflects self-financing due to the beneficial labor market consequences of the policy.

3.2 Welfare Optimal Policy

In this section we provide characteristics of optimal welfare policies.

Because of free entry, firms make no profit so the social welfare function consider welfare \( V_1 \) of \( N \) entrants in sector 1, welfare \( V_2 \) of \( (1 - N) \) entrants in sector 2 and the cost of entry in sector 1. Expected utility ex ante, prior to entry, equals average welfare ex post. An optimal policy maximizes social welfare \( V = \max \) \( t_1, r, b, t_1, T, m, t_2 \) \( NV_1 + (1 - N) V_2 - \int_0^N i(n) dn + \lambda T \) where \( \lambda \) is the Lagrange multiplier relating to the fiscal constraint \( T = 0 \). Even though the choice to join sector 1 or sector 2 is discrete for every worker, our assumption of uniform distribution of ability \( n \) and continuity in the firm creation effort \( i(n) \) leads to continuous variations in the endogenous sector 1 entry choice \( N \) and a social welfare function \( V \) which is continuous in all its arguments. This supports the existence of a solution to the social welfare maximization problem. Due to occupational choice, a variation of \( N \) yields \( dV = (V_1 - V_2 - i) dN = 0 \). Welfare maximization thus implies \( dV = Nd \ V_1 + (1 - N) dV_2 + \lambda dT = 0 \). Substituting (13) and (16) yields the optimality conditions, collected in the following result:

**Lemma 1. (Welfare optimality conditions):** the social welfare maximization policies satisfy:

\[
\begin{align*}
\frac{dV}{dt_1} &= - \left[ u_1' - \left( 1 - \tau N \eta_1 \right) \lambda \right] w_1 (1 - s) N = 0, \\
\frac{dV}{dt_1} &= - \left[ u_1' - \left( 1 - \tau N \eta_1 \right) \right] w_1 (1 - s) N = 0, \\
\frac{dV}{db} &= \left[ u_r' - \left( 1 - \tau N \eta_r + \tau E \varepsilon r \right) \lambda \right] (1 - e) sN = 0, \\
\frac{dV}{ds} &= - \left[ (1 - t_1 - \nabla \sigma) u_1' - \left( 1 - t_1 - \tau S \sigma - \tau N \eta_3 \right) \lambda \right] sN = 0, \\
\frac{dV}{dm} &= \left[ - \phi' \xi - \left( k - \tau N \eta m - \tau E \varepsilon m \right) \lambda \right] sN = 0, \\
\frac{dV}{dt_2} &= - \left[ u_2' - \left( 1 + \tau N \eta_2 \right) \lambda \right] w_2 (1 - N) = 0.
\end{align*}
\]

**Proof (lemma 1):** we provide the proof for the first condition and refer to appendix A for the rest, proved in similar fashion. Using (13), (16) and \( V_2 = u_2 = u \left( (1 - t_2) w_2 \right) \), we
have

\[
\frac{dV}{dt_1} = N \frac{dV_1}{dt_1} + (1 - N) \frac{dV_2}{dt_1} + \lambda \frac{dT}{dt_1}
\]

\[
= -N u'_r w_1 (1 - s) + 0 + \lambda \left(1 - \tau^N \eta_1\right) w_1 (1 - s) N
\]

\[
= - \left[u'_r - \left(1 - \tau^N \eta_1\right) \lambda N \right] w_1 (1 - s) N.
\]

QED.

Taking ratios to eliminate the shadow price \( \lambda \) leaves five optimality conditions which, together with the fiscal constraint, implicitly determine the optimal values of the six unknown policy variables. Several properties of optimal welfare policy can be derived. We collect the technical results in a proposition first and continue with an interpretation and discussion of the results.

**Proposition 1. (Optimal welfare characteristics):** the social welfare maximization policies have the following properties:

(a) there is limited unemployment insurance:

\[
\frac{u'_r}{u'_h} = \frac{1 - \tau^E \varepsilon_r}{1 + \tau^E \varepsilon_b} < 1 \quad u_r > u_h
\]

(b) there is overproportional earnings risk insurance:

\[
\frac{u'_r}{u'_1} = 1 - \tau^E \varepsilon_r < 1 \quad u_r > u_1
\]

(c) optimal tax redistribution is limited: assuming \( \tau^N > 0 \),

\[
\frac{u'_r}{u'_2} = 1 - \frac{\tau^N \eta_1}{1 - N} < 1 \quad u_1 > u_2
\]

(d) employment protection internalizes firing cost externalities:

\[
t_s = t_1 w_1 + [(1 - e) b - t_r w_r e] + km + \nabla
\]

(e) the welfare impact of firing depends on active labor market policies:

\[
\frac{d u^*}{d m} / \frac{d u_1}{d c_1} = -\phi' \zeta / u'_1 = k - \tau^E \varepsilon_m
\]

**Proof (proposition 1):** most of the results are direct derivations from lemma 1. We provide the proof for part (a). Proof for the other parts are similar and contained in appendix A.

Equating the second condition of the lemma \( \lambda = u'_r / \left(1 - \tau^N \eta_r - \tau^E \varepsilon_r\right) \) with the third condition \( \lambda = u'_h / \left(1 - \tau^N \eta_h + \tau^E \varepsilon_b\right) \) and using \( \eta_r = u'_r / N^r \) and \( \eta_h = u'_h / N^h \), one has

\[
u'_r \left(1 - \tau^N \eta_r / N^r + \tau^E \varepsilon_b\right) = u'_h \left(1 - \tau^N u'_r / N^h - \tau^E \varepsilon_r\right),
\]
that is
\[ u'_r \left( 1 + \tau^E \epsilon_b \right) = u'_h \left( 1 - \tau^E \epsilon_r \right). \]

Using the fact that elasticities \( \epsilon_j > 0 \) and \( \tau^E = t_r w_r + b > 0 \), it follows that
\[ \frac{u'_r}{u'_h} = \frac{1 - \tau^E \epsilon_r}{1 + \tau^E \epsilon_b} < 1. \]

The concavity of the utility function concludes.
QED.

We start the discussion with part (a) of the proposition, the fact that unemployment insurance (UI) is provided but in limited fashion. The left side of the condition
\[ \frac{u'_r}{u'_h} = \frac{1 - \tau^E \epsilon_r}{1 + \tau^E \epsilon_b} < 1 \]
reflects the marginal rate of substitution between consumption in the reemployed and the unemployed state (equal to \( -\epsilon \frac{u'_r}{1-\epsilon u'_h} \)) which indicates how much extra consumption an agent would need in the bad state \( (h) \) to compensate for one unit given up in the good state \( (r) \). The right side is proportional to the marginal rate of transformation (equal to the fraction times \( -\epsilon \frac{1}{1-\epsilon} \)) which states the rate at which government can shift consumption from the good to the bad state. If effort were inelastic (\( \epsilon \)-elasticities zero), optimal policy would implement full consumption smoothing between retraining and unemployment, \( u'_r = u'_h \). This is shown by the tangency point on the 45°-line in the right panel of Figure 2 where consumption in each state is denoted by \( c_j \), i.e. \( c_r = (1-t_r) w_r \). However, insurance diminishes incentives for job search and retraining. Providing insurance becomes more costly when it contributes to higher unemployment. It inflates social spending and simultaneously looses tax revenue when an agent switches from work to unemployment. The net effect on the fiscal budget is proportional to the participation tax \( \tau^E \) in both cases. Given that insurance becomes more costly, optimal policy advises limited insurance, leaving an income gap \( (1-t_r) w_r > b + h \) and \( u'_r < u'_h \).

We continue with the discussion of part (b) of the proposition. We have assumed that job separation in the volatile industry leads to a depreciation of sector specific skills. In consequence, the next best job opportunity pays only a lower wage, \( w_1 > w_r \). In other words, firing leads to an uninsured wage loss even if another job is found. Policy should thus provide to some extent insurance for earnings risk which is different from UI. The condition
\[ \frac{u'_r}{u'_1} = 1 - \tau^E \epsilon_r < 1 \]
reflects it. Without moral hazard in UI \( (\epsilon_r = 0) \), optimal policy aims at complete consumption smoothing between primary and reallocated employment, \( u'_1 = u'_r \), giving \( (1-t_1) w_1 = (1-t_r) w_r \). Full insurance requires a progressive rate structure, \( t_1 > t_r \).
If there is moral hazard, it becomes optimal, for given UI benefits, to strengthen incentives for retraining by shifting relatively more income towards reallocated employment. This calls for a higher tax rate on primary employment and an even more progressive tax system. With optimal policy, we thus have $(1 - t_1)w_1 < (1 - t_r)w_r$. The left panel of Figure 2 illustrates.\footnote{We assume that a job loss reduces utility so much that the participation constraint $u_1 > u^c$ is slack. With full consumption smoothing, $u_1 = u_e = u_0$, the utility loss, $u_1 - u^c = (1 - e)\chi + \phi \zeta (e)$ is positive if $\zeta$ and $\chi$ are positive, where $e$ is determined by $\chi = \phi \zeta' (e)$.}

Part (c) of the proposition characterizes the policy concern of redistribution between sector 1 workers who are in well paying, but risky jobs, and sector 2 workers where wages are safe but low (we assumed $w_1 > w_r > w_2 > h$). Optimal redistribution is characterized by

$$\frac{u_1'}{u_2'} = 1 - \frac{\tau^N \eta_1}{1 - N} < 1,$$

when the effective tax rate on sector 1 entry $\tau^N = t_1 w_1 - t_2 w_2 + s \tau^S > 0$, a condition which holds in realistic cases. Clearly, if the ability distribution were exogenous and a result of pure luck, $\eta_j = 0$, and investment in sector 1 firm creation were thus completely inelastic, optimal policy would also implement full redistribution, $u_1' = u_2'$ and $(1 - t_1)w_1 = (1 - t_2)w_2$. Due to $w_1 > w_2$, and ignoring for the moment other taxes and spending, the fiscal constraint $t_1 w_1 N + t_2 w_2 (1 - N) = 0$ requires a tax on sector 1 workers and a transfer $t_2$ to sector 2 workers. In general, we can write the fiscal budget as $t_2 w_2 = - \tau^N N$ where the effective tax $\tau^N$ on sector 1 workers finances a transfer $t_2 w_2 < 0$ to sector 2 workers. Imperfectly elastic firm creation investment prevents perfect income
smoothing since such redistribution diminishes incentives for initial firm creation investment and, thus, becomes increasingly costly. Hence, with positive $\eta$-elasticities, $u'_1 < u'_2$ and $(1 - t_1) w_1 > (1 - t_2) w_2$.

Part (d) of the proposition shows that government uses a firing tax to implement an optimal degree of job protection, given by the optimality condition

$$t_s = t_1 w_1 + [(1 - e) b - t_r w_r e] + km + \nabla.$$ 

The firing tax performs the same role as in Blanchard and Tirole (2008), even though the tax has new redistributive implications and additionally affects job creation and hiring. Its purpose is to internalize negative firing externalities. When dismissing a worker, the firm imposes an income equivalent utility loss $\nabla = (u_1 - u^e)/u'_1$ on that person. In addition, it creates a fiscal externality consisting of several components. First, there is one person less paying the tax $t_1$, and there is one person more who collects on average a net subsidy $(1 - e) b - t_r w_r e$, and there is extra spending $km$ per capita on ALMP. All these components might justify a substantial level of the firing tax.

Finally, part (e) of the proposition shows that active labor market policy (ALMP) can usefully complement other instruments and play a role in the welfare impact of firing. In raising policy amount $m$, the government spends a larger amount $km$ per capita of the $sN$ dismissed workers who are in need to be reallocated to another job. The left side of the optimality condition

$$\frac{d u^e}{dm} \frac{d u_1}{dc_1} = \frac{- \phi' \xi}{u'_1} = k - \tau^E \epsilon_m$$

states the marginal benefit of ALMP which raises expected utility $u^e$ of a fired relative to a retained worker. The relative welfare impact of firing thus depends on ALMP. The social cost of ALMP consists of the marginal resource cost $k$ per capita and is reduced by the budget savings $\tau^E \epsilon_m$ if ALMP puts a larger fraction of fired workers back to work. The elasticity $\epsilon_m$ measures how effective ALMP is to support job reallocation and boost reemployment among dismissed workers. The budget savings are proportional to the participation tax $\tau^E = t_r w_r + b$. Immervoll et al. (2007) found participation tax rates in Europe to vary mostly between 50 to 70% of gross wages, and up to 80% in Nordic countries. The upshot is that the participation tax and, in turn, the fiscal savings from ALMP are large in a generous welfare state with high benefits. These savings reduce the social cost of ALMP and lead to larger programs. ALMP programs thus become an essential ingredient of an advanced welfare state.

We finish this section with a special case. To put our analysis of flexicurity into perspective, we show that our framework is an extension of Blanchard and Tirole (2008), henceforth BT, and reproduce two of their central results. Excluding entry and job creation, we fix the mass of sector 1 workers at $N = 1$. Further, BT abstracted from job reallocation so that $e = 0$ and firing always results in unemployment. We then have
Lemma 2. (1 sector special case): assume that there is only one sector \((N = 1)\) with no job reallocation \((e = 0)\). Then full unemployment insurance is optimal, \(h + b = (1 - t_1)w_1\). Further, unemployment insurance is entirely financed by firing taxes \((t_1 = 0)\) when there is no job loss stigma \((\chi = 0)\). When there is job loss stigma \((\chi > 0)\), the firing tax is increased to reduce job separation, internalizing the firing externality and subsidizing employment \((t_1 < 0)\).

The lemma replicates propositions 1 and 2 of BT.

Proof (lemma 2): with the margins \(N = 1\) and \(e = 0\) fixed, the \(\varepsilon\)- and \(\eta\)-elasticities are all zero. Social welfare is \(V_1 = (1 - s)u_1 + s(u_h - \chi)\). The fiscal constraint reduces to \(T = t_1w_1(1 - s) + (t_s - b)s = t_1w_1 + s\tau^S = 0\), where \(\tau^S \equiv t_s - b - t_1w_1\) is the effective firing tax, or subsidy. The optimality conditions in (17) with respect to \(t_1, b\) and \(t_s\) are reduced to

\[ u_1' = \lambda = u_h', \quad \tau^S = \nabla. \] (18)

The optimal policy assures full consumption smoothing \((1 - t_1)w_1 = \lambda = h + b\). The stigma \(\chi\) of a job loss thus leads to a utility differential between work and unemployment equal to \(\nabla = (u_1 - u_h + \chi)/u_1' = \chi/u_1'\). Now suppose first that stigma is absent, so that \(\tau^S = \nabla = 0\). The fiscal constraint \(t_1w_1 = -s\tau^S\) then implies \(t_1 = 0\), and \(t_s = b\). Benefits are exclusively financed with a firing tax with no other tax on wages. Full consumption smoothing implies \(w_1 = h + b\). Substituting this into (2) leads to a firing threshold \(Ax_1 = w_1 - t_s = h + b - t_s = h\), i.e. \(Ax_1 = h\) as in Proposition 1 of BT. If there is a positive stigma, the effective firing tax \(\tau^S = \nabla\) is positive, implying \(t_s = b + t_1w_1 + \nabla\). The fiscal budget leads to a wage subsidy \(t_1w_1 = -s\nabla < 0\). Substituting \(t_s\) into the firing rule and noting full insurance yields \(Ax_1 = h - \nabla < h\). If there is stigma of job loss, the firing externality becomes larger. Optimal policy thus raises the firing tax to reduce job separation. Since this also depresses wages, workers are compensated by an employment subsidy \(t_1w_1 < 0\). These results replicate Proposition 2 of BT.

QED.

4 Policy Reform

In optimizing a social welfare function reflecting the allocative and distributive goals of public policy, the last section has provided criterias for an optimal flexicurity policy as part of the overall tax benefit schedule. The optimal policy with six policy instruments, \(t_1, t_r, b, t_s, m\) and \(t_2\), is only implicitly determined by the six conditions (17) together with the fiscal constraint. This section considers piecemeal reform. We aim to show how a small policy reform implementing steps towards an optimal policy leads to beneficial labor market and sectoral adjustment and, thereby, promises welfare gains. The first experiment starts out with an untaxed, free market equilibrium and introduces a flat rate UI scheme where all workers pay the same proportional contribution rate to finance
benefits of unemployed workers. This experiment resembles most current UI schemes which have largely flat contribution rates and often contain cross-subsidization between groups with different unemployment risks. In subsequent scenarios, we ‘improve’ on this scheme by adjusting the structure of tax rates and by complementing the tax benefit schedule with ALMP and with job protection by means of a firing tax. All scenarios must fulfill the fiscal constraint in (16), \( dT = 0 \), where effective tax rates are defined in (15). Inserting the welfare change of sector 1 employees in (13), social welfare changes by \( dV = NdV_1 + (1 - N) dV_2 \), or

\[
dV = -wresNdtr - w_1 (1 - s) Ndt - w_2 (1 - N) dt + u_r (1 - e) sNd - \phi wresNds - (1 - t - \nabla \sigma) u_1 sNdts.
\]

Flat Unemployment Insurance: Suppose the government finances benefits with a flat tax, \( t_1 = t_r = t_2 = t_s = m = 0 \), and abstains from any other labor market intervention, i.e. \( t_s = m = 0 \). Since there are gains from insurance, the policy must be welfare improving, at least if it is operated at a small scale. The following result confirms this expectation and characterizes employment and job creation effects:

**Proposition 2. (Introducing flat UI):** assume there is no government \( (t_1 = t_r = t_2 = t_s = b = m = 0) \) and unemployment insurance is introduced \( (db > 0) \), financed by a flat tax \( (dt_1 = dt_r = dt_2 = dt > 0) \). Then welfare \( V \) is improved, re-employment \( e \) is reduced, job creation \( N \) in the volatile sector and aggregate unemployed \( \delta \) both increase:

\[
\frac{dV}{db} > 0, \quad \frac{de}{db} < 0, \quad \frac{dN}{db} > 0, \quad \frac{d\delta}{db} > 0.
\]

**Proof (proposition 2):** in this experiment, the wage structure is exogenous to the policy change since the wage \( w_1 \) is fixed by \( ts = 0 \). Starting from a free market equilibrium with zero taxes implies effective tax rates \( \tau_j = 0 \). Financing small UI benefits \( db > 0 \) and satisfying the fiscal budget constraint in (16) requires a small increase in the uniform tax rate \( dt > 0 \) of

\[
\Gamma dt = (1 - e) sNd, \quad \Gamma \equiv w_1 (1 - s) N + w_r esN + w_2 (1 - N),
\]

where \( \Gamma \) is the tax base consisting of average wage earnings of individuals in different jobs. Since \( dt \) and \( db \) are positively correlated, one can prove the results differentiating by \( db \) or \( dt \). Substituting into the welfare differential and evaluating at \( t = b = 0 \) yields

\[
\frac{dV}{dt} \bigg|_{b=0} = \left[ 1 - \frac{w_1 w_2 (1 - N)}{\Gamma} - \frac{w_1 w_1 w_2 (1 - N)}{\Gamma} \right] u_h' > 0.
\]

The income distribution is characterized by \( w_1 > w_r > w_2 > h \), implying marginal rates of substitution all smaller than one. Since the weights of these ratios add up to unity, the square bracket is positive.
4. POLICY REFORM

The flat UI scheme, by assumption, does not make use of the firing tax, so that the wage and the separation rate in the volatile industry, \( w_1 \) and \( s \), remain constant. Substituting the first equation into (12) shows that the policy reform diminishes employment,

\[
\frac{de}{dt} = - \left[ ew_r \epsilon_r + \left( \frac{\Gamma}{sN} \right) \epsilon_h \right] < 0,
\]
as the square bracket only contains positive elements.

Substituting the fiscal balancing rule stated above into the entry response in (14) and noting the definition of \( \Gamma \) yields

\[
\frac{dN}{dt} = (\eta_h \eta_1) w_1 (1 - s) N + (\eta_h \eta_r) w_r e s N + (\eta_h \eta_2) w_2 (1 - N) > 0.
\]
The impact is unambiguously positive since \( w_1 > w_r > h \) implies \( u'_{1} < u'_{r} < u'_{h} \) and, therefore, \( \eta_1 < \eta_r < \eta_h \), when evaluated in the initial untaxed equilibrium.

Aggregate unemployment reflects job creation as well as firing and job search. The unemployment rate is \( \delta = (1 - e) s N \). Given a constant firing rate, the scenario changes unemployment by \( d\delta = -sNde + (1 - e) s dN > 0 \).

QED.

Offering small UI clearly raises welfare when private insurance markets are missing. Since tax distortions are initially small, the gains from insurance more than justify the tax cost. It is also intuitive that UI benefits and the contribution tax discourage retraining and search effort so that job reallocation slows down and unemployment among dismissed workers rises.

Further, it is interesting to note that the policy implicitly cross-subsidizes from sector 2 to sector 1 workers and, thus, encourages more people to enter the volatile industry. Entry results from occupational choice and is driven by \( V_1 - i(N) = V_2 \). Raising the tax rate clearly reduces utility from sector 2 work since the policy extracts tax from sector 2 workers without any compensation. Sector 1 workers also pay tax when employed, but are offered benefits when unemployed. In net terms, sector 1 workers benefit from cross-subsidization so that expected utility \( V_1 \) rises. Both effects stimulate investments in sector 1 firm creations and encourage entry.

Increase in aggregate unemployment is also intuitive. More entry exposes a larger fraction of the population to unemployment risk. Further, reallocation of dismissed sector 1 workers slows down and a larger share of them ends up unemployed. Both factors contribute to more unemployment. Since occupational choice determines the economy’s sectoral composition, the impact of entry on the unemployment rate also reflects structural change which shifts employment from stable to highly volatile sectors or, equivalently, from sectors with low to sectors with high unemployment incidence. Thereby, the aggregate unemployment rate is a function of the economy’s sectoral structure.

We now perform further policy experiments, assuming that a flat UI scheme is in place but no other tax or labor market policy is used and introducing policy instruments separately.
**Sectoral Redistribution:** One outcome of the introduction of flat UI is a relatively large participation tax on dismissed workers, $\tau^E = tw_r + b > 0$. In contrast, even though there is no statutory firing tax, $t_s = 0$, the effective tax on firing is negative, $\tau^S = [etw_r - (1 - e) b] - tw_1 < 0$, since UI among dismissed workers is cross-subsidized by other groups so that the square bracket is negative. If UI financing were strictly limited to dismissed workers, insurance would be actuarially fair, $etw_r = (1 - e) b$. Given a negative effective tax $\tau^S$, the flat UI scheme ends up subsidizing firing of workers in the volatile industry. Firing thus leads to a loss in the fiscal budget proportional to $\tau^S$. Finally, the policy cross-subsidizes from sector 2 to sector 1 workers and, thereby, encourages entry and job creation in sector 1. This is seen by the fiscal constraint $T = tw_2 + \tau^N N = 0$, which implies $\tau^N < 0$ when $t > 0$. With the effective tax rate on sector 1 entry $\tau^N = tw_1 - tw_2 + s \tau^S$ negative, the effective firing tax $\tau^S$ must actually be strongly negative. Cross-subsidization from sector 2, where wages are lower, to sector 1 may be incompatible with policy objectives.

It is indeed reasonable to expect that the flat tax rate structure to finance UI does not satisfy the distributional concerns in policy making. Given concave utility and the fact that earnings are lower in sector 2, a welfare based policy is more likely to call for redistribution from sector 1 to sector 2, rather than vice versa. Starting from the flat UI scheme, the next policy experiment raises the tax rate on sector 1 workers to cut the tax burden on low wage earners in sector 2. The following result characterizes welfare and employment effects:

**Proposition 3. (Adding sector redistribution):** assume policy as introduced in proposition 2 and further add redistribution from sector 1 to sector 2 ($dt_1 > 0 > dt_2$) such that the government budget remains balanced. Then the policy raises welfare $V$, leaves re-employment $e$ unaffected, discourages entry $N$ in the volatile sector 1 and reduces aggregate unemployment $\delta$:

$$
\frac{dV}{dt_2} < 0, \quad \frac{de}{dt_2} = 0, \quad \frac{dN}{dt_2} > 0, \quad \frac{d\delta}{dt_2} > 0.
$$

**Proof (proposition 3):** by (16), budget balance dictates a higher tax rate on sector 1 employees when $t_2$ is reduced,

$$
w_1 (1 - s) N dt_1 = -\frac{1 + \tau^N \eta_2}{1 - \tau^N \eta_1} w_2 (1 - N) dt_2.
$$

The flat UI scheme redistributes towards sector 1, $\tau^N < 0$, which is the ‘wrong’ direction. Evaluating (19) shows how redistribution towards sector 2 boosts welfare,

$$
dV = - \left[ \frac{u'_2}{u_1'} - \frac{1 + \tau^N \eta_2}{1 - \tau^N \eta_1} \right] u'_1 w_2 (1 - N) dt_2.
$$
Since wages satisfy $w_1 > w_2$, the square bracket is clearly positive. Hence, starting from a flat scheme and moving towards a more progressive rate structure ($dt_1 > dt_r = 0 > dt_2$) is welfare improving, $dV/dt_2 < 0$. Since this policy reform keeps not only $t_s = m = 0$ constant but also $t_r$ and $b$, it has no impact on the separation rate $s$ and on job search and labor reallocation $e$, so $de/dt_2 = 0$. The only effect is on entry and job creation $N$. Evaluating (14) subject to the budget balance condition yields

$$dN = \left[ \eta_2 + \eta_1 \frac{1 + \tau^N \eta_2}{1 - \tau^N \eta_1} \right] w_2 (1 - N) dt_2 < 0.$$ 

Both the lower tax rate on sector 2 and the higher rate on sector 1 employment discourage entry $N$ and employment in the volatile industry. Since neither the firing rate $s$ nor the rate $e$ of job reallocation are affected, the policy also squeezes aggregate unemployment, $d\delta = (1 - e) s dN < 0$.

QED.

When a smaller part of the population gets employed in the volatile sector, a smaller part is subject to unemployment risk. Structural change contributes to lower unemployment when sectors with high unemployment rates shrink and sectors with a low unemployment incidence expand.

**Subsidizing Job Reallocation:** While redistribution calls for a progressive tax structure ($t_1 > t_2$), the government might want to reduce the tax rate $t_r$ below $t_1$ by complementing the wage tax schedule with special tax credits or wage subsidies to reemployed workers in order to encourage retraining, thereby reducing the rate $t_r$ below $t_1$. In fact, the tax schedule might actually become non-monotonic with $t_1 > t_2 > t_r$ in spite of $w_1 > w_r > w_2$ as it clearly does in the optimal policy scenarios of the next section.

The rationale of this policy rests on the following arguments. Wage taxes and UI discourage retraining and job search of fired workers and reduce the transition rate $e$ into alternative employment. When switching from unemployment into a job, individuals face a very high participation tax $\tau^E = t_r w_r + b$, as households leaving unemployment need to pay tax $t_r w_r$ and give up benefits $b$. Bringing down the participation tax and thereby encouraging retraining and job search helps to speed up job reallocation. Since high benefits $b$ are needed to provide insurance, the only way to do so is the reduce the tax rate $t_r$. This motivates a reform shifting the tax burden from reemployed workers ($dt_r < 0$), the middle income wage earners in our model, towards the top (and bottom) income class ($dt_1 > 0$). The next result defines welfare and employment effects:

**Proposition 4. (Adding job reallocation subsidies):** assume policy as introduced in proposition 2 and further add redistribution (wage subsidies) from sector 1 to re-employed workers in sector 2 ($dt_1 > 0 > dt_r$) such that government budget is balanced. Then the policy raises welfare $V$, increases re-employment $e$, supports entry $N$ in the volatile
sector 1 but has ambiguous effects on aggregate unemployment $\delta$:

$$\frac{dV}{dt_r} < 0, \quad \frac{de}{dt_r} < 0, \quad \frac{dN}{dt_r} < 0, \quad \frac{d\delta}{dt_r} = ?.$$  

Proof (proposition 4): to balance the government budget with the reform shifting the tax burden from reemployed workers, the middle income wage earners in our model, towards the top income class, one uses:

$$w_1 (1 - s) N dt_1 = -\frac{1 - \tau^N \eta_r - \tau^E \varepsilon_r}{1 - \tau^N \eta_1} w_{res} N dt_r.$$  

Starting from the flat tax benefit schedule, the policy impacts welfare by

$$dV = -\left[ u_r' - \frac{1 - \tau^N \eta_r - \tau^E \varepsilon_r}{1 - \tau^N \eta_1} \right] u_1' w_r \varepsilon s N dt_r.$$  

Since $u_r' > u_1'$ and $\eta_r = u_r' / N_1' > \eta_1 = u_1' / N_1'$ in the initial equilibrium with a flat scheme, the square bracket is clearly positive. Hence, raising $t_1$ to finance a budget neutral tax cut for reemployed workers, $dt_r < 0$, boosts welfare: $dV / dt_r < 0$. Since firing depends neither on $t_1$ nor on $t_r$, the separation rate $s$ is unchanged. However, the cut in the participation tax $\tau^E = t_r w_r + b$ stimulates job search and reemployment of dismissed workers by $de = -\varepsilon_r w_r edt_r$, so $de / dt_r < 0$. Using the fiscal balance condition above into (14) and noting $\eta_r / \eta_1 = u_r' / u_1'$ yields

$$dN = -\left[ u_r' - \frac{1 - \tau^N \eta_r - \tau^E \varepsilon_r}{1 - \tau^N \eta_1} \right] \eta_1 w_r \varepsilon s N dt_r.$$  

By the same argument as before, the lower tax on reemployed workers, $dt_r < 0 < dt_1$, boosts entry and employment in the volatile industry, $dN / dt_r < 0$. The effect on aggregate unemployment $\delta = (1 - e) s N$ is ambiguous since $1 - e$ falls while entry $N$ rises.

QED.

**Job Protection:** The next reform introduces job protection to complement the flat UI scheme. Since the firing tax $t_s$ is an extra cost, sector 1 firms compete down the wage $w_1$ to break even in zero profit equilibrium. To compensate workers, the scenario cuts the tax rate $t_1$. The following proposition characterizes welfare and employment effects:

**Proposition 5. (Adding employment protection):** assume policy as introduced in proposition 2 and further add job protection financing labor income tax reduction in the volatile sector ($dt_s > 0 > dt_1$) to maintain government budget balance. Then the policy raises welfare $V$, leaves re-employment $e$ unchanged, supports entry $N$ in the volatile sector 1 but has ambiguous effects on aggregate unemployment $\delta$:

$$\frac{dV}{dt_s} > 0, \quad \frac{de}{dt_s} = 0, \quad \frac{dN}{dt_s} > 0, \quad \frac{d\delta}{dt_s} = ?.$$
4. POLICY REFORM

Proof (proposition 5): the increase in firing tax $t_s$ is associated with a decrease in sector 1 tax $t_1$ to maintain fiscal budget balance,

$$w_1 (1-s) N dt_1 = - \frac{1 - t - \tau^S \sigma - \tau^N \eta_s}{1 - \tau^N \eta_1} s N dt_s.$$ 

Substituting into the welfare change stated in (19) and using $\eta_s = (1 - t_1 - \nabla \sigma) \eta_1$ yields

$$dV = u'_1 \frac{\nabla - \tau^S}{1 - \tau^N \eta_1} s N dt_s.$$ 

Job protection reduces turnover, $ds = -\sigma s dt_s$, so that workers are exposed less frequently to the utility loss $\nabla u'_1 = u_1 - u^e$. The firing tax, however, leads to an a priori ambiguous effect on disposable income of sector 1 employees. Since the flat UI scheme subsidizes firing, i.e. the effective tax rate on firing is $\tau^S < 0$, the reduced job separation yields a fiscal dividend in proportion to $\tau^S$ which allows a relatively larger tax cut $t_1$. Disposable earnings and welfare increase on this account. However, the firing tax $t_s$ also forces firms to cut the gross wage $w_1$. Depending on the size of the tax cut $t_1$ relative to the wage reduction in $w_1$, the net of tax wage may rise or fall. Using the fiscal balance condition and (5), and making use of $\eta_s = (1 - t_1 - \nabla \sigma) \eta_1$, yields

$$d [(1 - t_1) w_1] = \frac{\nabla \tau^N \eta_1 - \tau^S \sigma s}{1 - \tau^N \eta_1} s N dt_s.$$ 

Sector 1 employees benefit from job protection since $\tau^S < 0$ at the outset. Despite of a lower gross wage, the compensating tax cut more than offsets this income loss, i.e. $u_1$ rises while $u^e$ remains constant. The introduction of a firing tax clearly boosts welfare, $dV/dt_s > 0$.

Entering sector 1 thus promises higher expected welfare which induces more workers to invest in firm creation and allows the volatile industry to expand employment. Evaluating (14) and using the same steps as before results in

$$dN = \frac{\nabla - \tau^S \sigma}{1 - \tau^N \eta_1} \eta_1 s N dt_s,$$ 

thus $dN/dt_s > 0$. Shifting labor to the volatile sector inflates unemployment $\delta = (1 - e) s N$ since unemployment is concentrated there. Hence, joblessness rises with the expansion of the volatile industry $N$ while a lower separation rate $s$ reduces unemployment. The reemployment rate $e$ is constant in this scenario $(de/dt_s = 0)$ since expected utility $u^e$ is unchanged. As a consequence, unemployment changes by $d\delta = (1 - e)[s dN + N ds]$. Substituting the change in entry $dN$ and job separation $ds$ yields an ambiguous net effect on aggregate unemployment $(d\delta/dt_s)$, given by

$$\frac{1}{(1 - e) \sigma s N dt_s} d\delta = - \left[ 1 - \frac{\nabla - \tau^S}{1 - \tau^N \eta_1} \eta_1 \right] = - \frac{1 - (\nabla s + t w_1 - t w_2) \eta_1}{1 - \tau^N \eta_1}, \quad (20)$$

where the second equality uses $\tau^N = t w_1 - t w_2 + s \tau^S$. 

QED.

The proposition shows that the policy effect on aggregate unemployment is ambiguous. To identify the conditions favoring a rise or fall in unemployment, one may interpret a large utility differential $\nabla = (u_1 - u^e) / u_1'$ as a high degree of sector specificity of skills. High sector specificity means that a worker suffers a large loss in wage and utility $\nabla$ when loosing her job. Further, a high separation rate $s$ stands for short employment tenure and high job-turnover, characterizing a very volatile industry. Finally, a large value of $\eta_1$ means that entry $N$ into sector 1 is very elastic, since $dN/dt_1 = -\eta_1 w_1 (1 - s) N dt_1$. Starting with a flat UI scheme, a small firing tax changes the unemployment rate as in the last term in (20). If sector 1 is characterized by a large degree of sector specificity of skills ($\nabla$ large), high job turnover (large $s$), and elastic labor supply (high $\eta_1$, reflecting elastic entry), a firing tax combined with a budget neutral wage subsidy ($dt_s > 0 > dt_1$) could actually raise unemployment. The reason is that all these conditions favor entry and a large expansion of the volatile industry where the sectoral unemployment rate is high, while the stable sector with a low (zero) unemployment rate strongly shrinks. The resulting increase in the unemployment rate could dominate the reduction in unemployment on account of a lower separation rate. However, this is not possible when the firing tax is close to its optimal value, implying $\nabla = \tau^S$, and is further increased. In this case, the unemployment rate unambiguously falls, as shown in (20).

**Active Labor Market Policy:** Finally, we analyze the consequences of complementing the flat UI scheme by introducing ALMP to facilitate retraining, search and job reallocation. Again, we raise the wage tax of workers in the volatile industry since the policy benefits them by improving their labor market prospects if they get fired. The following result analyzes welfare and employment effects:

**Proposition 6. (Adding active labor market policy):** assume policy as introduced in proposition 2 and further add active labor market policies ($dm > 0$) financed by a higher labor income tax in the volatile sector ($dt_1 > 0$) to maintain government budget balance. Then the policy has ambiguous effects on welfare $V$, entry $N$ in the volatile sector 1 and aggregate unemployment $\delta$, but supports re-employment $e$. More precisely:

$$\text{sign} \left( \frac{dV}{dm} \right) = \text{sign} \left( -\phi' \frac{\zeta}{u_1'} - (k - \tau^E e_m) \right), \quad \frac{de}{dm} > 0, \quad \frac{dN}{dt_s} \bigg|_{dt_s > 0} > 0, \quad \frac{d\delta}{dm} \bigg|_{\frac{dv}{dm} = 0} < 0.$$

**Proof (proposition 6):** spending more an ALMP by raising $m$ thus requires a higher tax $t_1$ on sector 1 employees given by

$$w_1 (1 - s) N dt_1 = \frac{k - \tau^N \eta_m - \tau^E e_m}{1 - \tau^N \eta_1} s N dm.$$
Welfare rises if the gains from better labor market prospects after firing dominate the welfare loss from a higher tax on sector 1 employment. Using \( \eta_m = -\phi' \zeta \eta_1 / u_1' \) yields a welfare change that is related to the optimality condition in part (e) of proposition 1,

\[
dV = \left[ \frac{-\phi' \zeta}{u_1'} - (k - \tau^E \epsilon_m) \right] \frac{u_1'}{1 - \tau^N \eta_1} sNdm.
\]

The sign of \( dV / dm \) is then given by the sign of \( -\phi' \zeta / u_1' - (k - \tau^E \epsilon_m) \). ALMP boosts job search by \( de = \epsilon_m dm \). Given a constant firing tax \( t_s \), the separation rate \( s \) remains unchanged. If the policy boosts welfare, it also encourages entry \( N \). Indeed, imposing budget balance and using \( \eta_m = -\phi' \zeta \eta_1 / u_1' \) yields an entry response in (14) equal to

\[
dN = \left[ \frac{-\phi' \zeta}{u_1'} - (k - \tau^E \epsilon_m) \right] \frac{\eta_1}{1 - \tau^N \eta_1} sNdm.
\]

ALMP cannot influence unemployment via the firing channel since \( s \) is fixed in this scenario. On the positive side, ALMP clearly speeds up job reallocation \( e \) and reduces the unemployment risk after a job loss. However, by encouraging entry \( N \) into the volatile sector, it also raises the absolute levels of job separation which adds to unemployment \( \delta = (1 - e) sN \). The net effect on aggregate unemployment \( \delta \) is a priori ambiguous. However, if ALMP is at its optimal level, \( -\phi' \zeta / u_1' - (k - \tau^E \epsilon_m) = 0 \) and ALMP does not affect entry and sectoral composition at the margin. In this case, ALMP reduces unemployment by stimulating job search.

QED.

The proposition shows that the introduction of active labor market policies has an ambiguous effect on welfare. Whether ALMP is welfare improving depends on its potency to raise expected utility of fired workers, \( du^E / dm = -\phi' \zeta > 0 \), relative to its cost, \( k - \tau^E \epsilon_m \). Obviously, if it is too costly \( (k \) high), the policy reduces welfare and is not advised. However, if high UI benefits \( b \) are offered to insure workers, the participation tax \( \tau^E = t_s w_r + b \) is high and the policy yields large fiscal gains. In bringing more people back to work, the policy boosts tax revenue and contains social spending. Hence, the net fiscal cost of ALMP is greatly reduced in the presence of a large welfare state. This is even more the case if the policy is very effective in reallocating workers \( (high e) \), as indicated by a large elasticity \( \epsilon_m = de / dm \). Given that ALMP is sufficiently productive, it becomes an essential ingredient of the welfare state.

ALMP are thus more interesting in countries with large welfare states, where unemployment benefits levels are high, and the cost of the ALMP are low. While the second feature is intuitive, the first feature may not be immediately. The reason is that every time an unemployed finds a new job the heavy tax burden of unemployment insurance is reduced, so it is worth investing more into job search assistance. This could explain why flexicurity measures were first introduced by Northern European countries in the 1990’s, who started with high unemployment benefits levels and subsequently introduced ALMP.
Summing up results in propositions 2, 3, 4, 5 and 6, the piecemeal reforms which introduced policy instruments separately, starting from no policy, increases welfare in most cases. We have proved that a measured introduction of unemployment insurance, partial redistribution from high paying to low paying sectors, wage subsidies to encourage re-employment of fired workers and employment protection all improve welfare. Active labor market policies have an ambiguous effect on welfare, which depends on their cost and the level of unemployment benefits: when costs are low and benefits high, active policies improve welfare. Impacts on aggregate unemployment are most of the time ambiguous. Only redistribution reduces unemployment, while unemployment insurance increases it.

5 Comparative Statics of Optimal Policy

This section numerically illustrates our results, in the first part, and shows how optimal flexicurity changes with risk-aversion, volatility and behavioral elasticities, in the second part. In all cases, we recalibrate the model to yield the same data. Details on the calibration are contained in appendix B.

5.1 Baseline

Column ‘Flat’ in Table 1 reports equilibrium values when policy is not optimized and neither a firing tax nor ALMP are used. The tax schedule is flat with a rate of 18%. Revenues finance public consumption equal to 15% of GDP, and UI benefits. Public goods per capita are fixed and, thus, do not affect welfare analysis. Initially, 60% of the labor force is attracted by higher wages and ready to become employed in the volatile sector where the separation rate is 25%. After separation, 60% get reemployed in the safe sector, the rest ends up unemployed. Unemployment amounts to 6%, reflecting job creation (entry), job destruction (firing) and the frictions in job reallocation (search). Ultimately, only 45% of the workforce remain employed in the volatile industry and 49% end up in the safe sector. The GDP share of the volatile sector is about 55%.

Column ‘Base’ reports the consequences of fully optimizing flexicurity and tax policy. Note first the effective tax rates in the status quo as defined in (15). The effective tax on job search is a high .78 which amounts to 68% of a retrained worker’s wage ($\tau_E/w_r = .68$, which is the sum of the tax rate and UI replacement rate). Such high values are consistent with the calculations of participation tax rates in Europe by Immervoll et al. (2007). More surprisingly, the status quo involves a high firing subsidy equal to $\tau_S = -.35$ which amounts to about 25% ($= \tau_S/w_1$) of a worker’s salary. This subsidy consists of the tax loss from each person that is fired ($t_1w_1 = .25$) plus the net budget cost of a fired worker (equal to $et,w_r - (1 - e) b = -.1$). Firing imposes a substantial

---

12In a dynamic model, the separation rate is the inverse of job tenure, $1/s = 4$ years in our case.
5. **COMPARATIVE STATICS OF OPTIMAL POLICY**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Flat</th>
<th>Base</th>
<th>RA</th>
<th>JS</th>
<th>Vol</th>
<th>Entry</th>
<th>-Alm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tax sector 1</td>
<td>0.180</td>
<td>0.149</td>
<td>0.143</td>
<td>0.150</td>
<td>0.138</td>
<td>0.165</td>
<td>0.146</td>
</tr>
<tr>
<td>Tax sector 2</td>
<td>0.180</td>
<td>0.109</td>
<td>0.125</td>
<td>0.110</td>
<td>0.112</td>
<td>0.084</td>
<td>0.109</td>
</tr>
<tr>
<td>Tax retrained</td>
<td>0.180</td>
<td>0.020</td>
<td>0.027</td>
<td>0.022</td>
<td>0.024</td>
<td>0.037</td>
<td>-0.005</td>
</tr>
<tr>
<td>Replacement rate UI</td>
<td>0.500</td>
<td>0.418</td>
<td>0.569</td>
<td>0.579</td>
<td>0.414</td>
<td>0.401</td>
<td>0.525</td>
</tr>
<tr>
<td>Firing tax</td>
<td>0.000</td>
<td>0.885</td>
<td>0.946</td>
<td>0.917</td>
<td>0.861</td>
<td>0.873</td>
<td>0.918</td>
</tr>
<tr>
<td>Search assistance</td>
<td>0.000</td>
<td>0.394</td>
<td>0.264</td>
<td>0.126</td>
<td>0.392</td>
<td>0.383</td>
<td>0.000</td>
</tr>
<tr>
<td>Eff. tax search</td>
<td>0.783</td>
<td>0.503</td>
<td>0.686</td>
<td>0.691</td>
<td>0.503</td>
<td>0.503</td>
<td>0.598</td>
</tr>
<tr>
<td>Eff. tax firing</td>
<td>-0.358</td>
<td>0.415</td>
<td>0.433</td>
<td>0.433</td>
<td>0.414</td>
<td>0.409</td>
<td>0.516</td>
</tr>
<tr>
<td>Eff. tax entry</td>
<td>-0.017</td>
<td>0.105</td>
<td>0.078</td>
<td>0.104</td>
<td>0.102</td>
<td>0.150</td>
<td>0.106</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Economic impact:</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Separation rate</td>
<td>0.250</td>
<td>0.074</td>
<td>0.063</td>
<td>0.068</td>
<td>0.114</td>
<td>0.076</td>
<td>0.068</td>
</tr>
<tr>
<td>Reemployment rate</td>
<td>0.600</td>
<td>0.778</td>
<td>0.653</td>
<td>0.620</td>
<td>0.778</td>
<td>0.779</td>
<td>0.636</td>
</tr>
<tr>
<td>Entry rate</td>
<td>0.600</td>
<td>0.579</td>
<td>0.580</td>
<td>0.579</td>
<td>0.570</td>
<td>0.575</td>
<td>0.578</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>0.060</td>
<td>0.010</td>
<td>0.013</td>
<td>0.015</td>
<td>0.014</td>
<td>0.010</td>
<td>0.014</td>
</tr>
<tr>
<td>Reallocation rate</td>
<td>0.150</td>
<td>0.058</td>
<td>0.041</td>
<td>0.042</td>
<td>0.089</td>
<td>0.059</td>
<td>0.043</td>
</tr>
<tr>
<td>Wage rate sector 1</td>
<td>1.400</td>
<td>1.229</td>
<td>1.224</td>
<td>1.226</td>
<td>1.208</td>
<td>1.230</td>
<td>1.226</td>
</tr>
<tr>
<td>Sector 1, % of GDP</td>
<td>0.556</td>
<td>0.603</td>
<td>0.610</td>
<td>0.608</td>
<td>0.577</td>
<td>0.598</td>
<td>0.607</td>
</tr>
<tr>
<td>GDP, (%)</td>
<td>0.000</td>
<td>2.008</td>
<td>1.226</td>
<td>1.200</td>
<td>1.830</td>
<td>1.969</td>
<td>1.224</td>
</tr>
<tr>
<td>Welfare, *)</td>
<td>0.000</td>
<td>6.492</td>
<td>7.365</td>
<td>6.354</td>
<td>5.396</td>
<td>6.649</td>
<td>6.359</td>
</tr>
</tbody>
</table>

*Legend:* All values in absolute terms, (%) change in percent. *) Welfare change in percent of initial GDP, 100 * (V - V₀)/X₀. (FLAT): flat tax rate and no welfare policy. (BASE): policy with base line parameters $\rho = 1.75$, $\epsilon = .4$, $\text{var}(x) = 3$ and $\eta = .8$. (RA): high value of risk-aversion $\rho = 3$. (JS): low elasticity of job search $\epsilon = .2$. (VOL): high volatility $\text{var}(x) = 5$. (ENTRY): low elasticity of entry $\eta = .4$. (-ALM): no active labor market policy $m = 0$.

Table 1: Comparative Statics of Optimal Welfare Policy

...net cost on the fiscal budget. Finally, the initial flat tax equilibrium also involves a small entry subsidy $\tau^\text{N}$. Although the tax liability of sector 1 workers exceeds that in sector 2 by $t \cdot (w_1 - w_2)$, this net entry tax is more than offset by the firing subsidy $s\tau^S$ that sector 1 workers receive whenever they lose their job.

With baseline parameter values, moving to a fully optimized policy leads to several adjustments of instruments. First, reflecting distributional concerns, the government moves to a progressive wage tax schedule by raising the tax rate on high sector 1 earnings relative to the tax rate on lower sector 2 earnings, $t_1 > t_2$. Second, it strongly reduces the tax rate $t_2$ to encourage job search and retraining after separation from sector 1 employment. The difference $t_1 - t_r$ may be interpreted as a wage subsidy to facilitate job reallocation. Third, and for the same reason, the government reduces unemployment benefits of dismissed workers, leading to a lower replacement rate of .41 instead of .5 initially. This means that the initial replacement rate was too high, given the degree of risk-aversion. As a result, the effective tax on job search is much reduced from .78 to .5. Fourth, the government imposes a large firing tax to internalize the negative firing externalities which turns the effective subsidy of $-.36$ into an effective tax on firing...
equal to .4. In consequence, the separation rate is strongly reduced from 25 to 7% of initially hired workers which, in turn, reduces the reallocation rate $e$ despite of a higher reemployment rate $e$ among fired workers. The firing tax also raises significant public revenues and contributes to an overall lower tax burden. Fifth, the government also introduces an ALMP program to reduce private search costs and facilitate retraining, and spends .7 percent of GDP for this purpose. Together with the fiscal incentives provided by the lower participation tax $\tau^E$, the policy boosts the reemployment rate $e$ (from 60 to 78%) as a result of successful job search.

Implementing the optimal policy actually raises the entry tax. The definition of the effective entry tax $\tau^N = t_1 w_1 - t_2 w_2 + s \tau^S$ shows that this partly results from moving to a progressive tax structure reflecting distributional concerns so that sector 1 work is subject to an additional tax liability $t_1 w_1 - t_2 w_2$. Another reason is the Pigovian firing tax which leads to a reduction in expected private earnings in sector 1 equal to $s \tau^S$. The firing tax thereby reduces job creation and, together with the progressive wage tax, discourages employment in the volatile sector. Entry is further discouraged by the fact that the firing tax is shifted to workers by squeezing the gross wage from 1.4 to 1.23. The entry rate declines from 60 to 58%. All three behavioral responses taken together lead to a large reduction of the frictional unemployment rate from 6 to 1 percent, down by 5 percentage points. The decline in unemployment not only reflects a lower separation and a larger reemployment rate. Lower entry means a reallocation away from sectors with above average to sectors with below average (zero) unemployment. At the same time, the baseline scenario shows not only an increase in aggregate GDP (plus 2%) but also a larger GDP share of the volatile sector (up by 4 percentage points from 56 to 60% of GDP). Although the policy reduces entry, any given hiring results in much more employment on account of a lower separation rate.

Finally, moving to an optimal flexicurity policy yields substantial welfare gains, equal to 6.5% of GDP in the baseline case. Sector 2 workers benefit from a much lower tax. Expected utility $V_1$ of sector 1 work rises too but ex post utilities change in intricate ways. Workers who remain employed in sector 1, lose despite of a lower tax rate since the firing tax on firms strongly reduces the gross wage $w_1$, see (5). Workers reemployed after separation strongly benefit from a reduction in the net tax $t_r$. Separated workers who remain unemployed, get lower benefits and are worse off, but there are fewer since unemployment is strongly reduced. Despite these offsetting changes, expected welfare after separation clearly rises, also because job search assistance reduces search costs and boosts welfare. In general, the total welfare gain results from reducing excessive UI

---

13Our welfare change measure is defined as the percentage variation of initial GDP, $100 \times (V - V_0)/X_0$. Consistent with one practice in the quantitative international economics literature (such as Heijdra & all, 2002 and Brochner & all, 2007), we scale down welfare variations between the initial 'Flat' level $V_0$ and the optimal 'Base' level $V$ by the initial economic size (GDP) of the country $X_0$, to remove scale effects. Although population is constant and there is no growth in our static model, the measure we use allows for an easier comparison with this literature. We do not use a measure based on either equivalent or compensating variations because of the absence of a representative agent in the model.
to a lower optimal level;\footnote{Obviously, if there were no UI at the outset, there would be strong gains from insurance in a setting with missing private UI markets.} introducing a Pigovian tax to internalize firing externalities; providing job search assistance which yields welfare gains like a public good; and gains from redistribution towards sector 2 workers who enjoy lower utility at the outset.

5.2 Sensitivity analysis

The last five columns of Table 1 show how optimal policy changes with structural characteristics and should be compared to column ‘Base’. Column ‘RA’ refers to an economy where individuals are more risk-averse. Obviously, a larger degree of risk-aversion calls for more consumption smoothing. Starting out from the equilibrium given in column ‘Flat’, the immediate consequence is a larger replacement rate for UI. To a large part, the fiscal cost is paid for by raising the firing tax because firing now causes a larger externality. More generous UI weakens job search incentives. Interestingly, the optimal policy does not imply larger wage subsidies of reemployed workers to offset these incentives. In fact, the tax rate $t_r$ slightly rises. Compared to the baseline, the optimal policy accepts a larger effective tax rate on job search when risk-aversion is high. Instead, a rise in unemployment is largely prevented by containing job separation. Compared to the base case, the unemployment rate is up from 1 to only 1.3%.

Interestingly, ALMP is significantly scaled back. The intuition follows from the fact that the firing tax is shifted to sector 1 workers by lowering the gross wage, see (5), leading to lower utility of retained workers while the welfare of a fired worker rises both with UI benefits and search assistance, see (12). In trading off utility of retained and fired workers, the government thus cuts back on job search assistance when higher risk-aversion calls for more generous UI benefits to smooth consumption more evenly across states. Optimal policy also shifts towards a less progressive tax schedule, i.e. $t_1$ is reduced and $t_2$ increased, compared to the base case. Since employment in sector 1 is risky and safe in sector 2, intuition suggests that higher risk-aversion requires a larger risk-premium on sector 1 income. All adjustments result in a lower entry tax $\tau_N$, leading to a slight and rather insignificant increase in entry. Overall, Table 1 shows a lower GDP gain from implementing the optimal policy which mainly reflects the reduced effectiveness in job reallocation. Nevertheless, the welfare gains of moving to an optimal policy are larger, reflecting the larger gains from insurance.

Column ‘JS’ differs from the baseline only in that job search is less elastic. Optimal policy adjusts by raising UI benefits, leading to a replacement rate of 58 instead of 42% and to a higher effective tax on job search. Since a higher replacement rate raises the firing externality, the government imposes an even higher firing tax and, to balance utility of retained and fired workers, cuts back on job search assistance. The reemployment rate increases much less compared to the baseline while the separation rate declines only to
a moderate extent so that the reallocation rate falls. In consequence, the optimal policy results in a smaller GDP gain of 1.2 instead of 2%. Other changes are negligible.

Keeping other parameters at baseline values, column ‘Vol’ raises the variance of productivity shocks in sector 1 from 3 to 5. The average productivity is increased in a compensating way in order to keep the firing threshold and, thus, the separation and wage rates constant, leaving the flat tax equilibrium unaffected. Table 1 shows that the optimal policy basically remains unchanged. However, since a higher volatility reduces the firing elasticity (the larger dispersion of the uniform distribution implies a smaller density $g(x_1)$ in equation 5), the firing tax increases the separation rate to some extent, to 11% up from 7% in the baseline case. Since the firing tax is paid more often, increasing costs and forcing a somewhat larger reduction in the wage rate in zero profit equilibrium. Therefore, entry into sector 1 is reduced by about 1 percentage point. The higher firing rate magnifies and the lower entry rate shrinks unemployment. On net, unemployment is up by roughly half a percentage point. A higher volatility dampens the GDP and welfare gains from implementing the optimal policy package.

Column ‘Entry’ considers an economy with a lower entry elasticity. Compared to the baseline case, Table 1 indicates that less elastic entry facilitates redistribution between workers in volatile, high paying jobs, and safe jobs with moderate salaries. Optimal tax rates $t_1$ and $t_r$ are up by about 1.5 percentage points while the tax $t_2$ on sector 2 workers is reduced from 10.9 to 8.4%. The government thus adopts a more progressive tax structure (with $t_r$ being interpreted as $t_1$ minus a wage subsidy). The remaining parts of the optimal policy package as well as effective tax rates on search and firing essentially remain unchanged. Reflecting the larger redistribution, the effective entry tax rises from 10.5 to 15% which slightly impairs entry. The larger scope for redistribution also rises the welfare gain from moving towards an optimal policy.

The last column ‘-Alm’ repeats the baseline policy scenario except that the use of job search assistance is artificially excluded ($m = 0$). Obviously, this leads an optimizing government to adjust the remaining policy instruments. Instead of raising welfare of fired workers by reducing their effort cost of job search, the government offers higher unemployment benefits and reduces the wage subsidy, i.e. the tax $t_r$ of the re-employed. Since this adjustment inflates the firing externality, the optimal firing tax is increased as well. Effective tax rates on job search and firing are up while the entry tax is essentially unchanged. Compared to the base case, job separation is further reduced and reemployment less frequent. Excluding ALMP thus dampens the GDP gain, mainly due to weaker search incentives and moderately higher unemployment.
6 Conclusions

The process of creative destruction in an advanced economy leads to a constant reallocation of labor and creates frictional unemployment. The experimentation with new production techniques, high job turnover, the riskiness of employment relationships and, therefore, the incidence of unemployment varies across sectors. One may think of highly volatile innovative industries compared to more stable traditional sectors. Technological progress and exposure to international competition may be relatively more concentrated in highly productive manufacturing and internationally traded services while non-traded sectors and services are less productive but also less volatile with lower sectoral unemployment rates. If high productivity and skill intensity are correlated with volatility, then unemployment, compared to earlier times, should differ less along the high- and low-skilled dimension and more between employment in volatile compared to less volatile sectors.

More turbulent economic times, characterized by increasing globalization, rapid technological progress and ongoing restructuring, create pressure on welfare states. It is often informally argued that the flexicurity model might be a better solution in more turbulent times. Flexicurity is usually defined by three pillars: (i) flexibility in firing if jobs turn out unproductive and labor would be better used somewhere else; (ii) social insurance to protect against the income risk of job separation; and (iii) active labor market policy to facilitate labor reallocation to new employment opportunities. This paper proposes a two sector model to analyze flexicurity where one sector is more productive but highly volatile while the other offers stable employment at lower wages. In consequence, unemployment is concentrated in the volatile industry. Structural change happens on two margins. The first is firm creation investment and entry of workers into the volatile industry rather than the safe sector. Retraining of workers fired in the volatile industry for alternative employment in the safe sector is a second channel for structural change.

Within this framework, we have characterized optimal welfare policies combining the design of the wage tax schedule with social insurance, job protection and activation measures. We emphasize five results. First, tax progressivity is motivated by redistribution from workers in highly paid sector 1 jobs to those with lower wages in sector 2. To maintain incentive to join the risky sector 1 however, there is no full redistribution. Second, the tax schedule is complemented by special tax credits or subsidies to encourage retraining and search for alternative employment. Third, public unemployment insurance is necessary due to missing private markets. Optimal insurance is however partial, to encourage job search. Fourth, negative firingexternalityscreate a need for an optimal degree of job protection. And fifth, active labor market policy is an essential ingredient of a large welfare state if its costs are low, it is sufficiently productive in stimulating reemployment and raising individual welfare by reducing private search costs. Most policies have an ambiguous impact on aggregate unemployment, but tax progressivity reduce it, lower taxes in sector 2 boosting its attractivity for job seekers.
Appendices
A. Proof details

**Proof (lemma 1):** using (13), (16) and $V_2 = u_2 = u((1 - t_2)w_2)$, we have

\[
\frac{dV}{dt_1} = N \frac{dV_1}{dt_1} + (1 - N) \frac{dV_2}{dt_1} + \lambda \frac{dT}{dt_1}
\]

\[
= -Nu'_1 w_1 (1 - s) + 0 + \lambda \left(1 - \tau^N \eta_1\right) w_1 (1 - s) N
\]

\[
= - \left[u'_1 - (1 - \tau^N \eta_1) \lambda\right] w_1 (1 - s) N
\]

and

\[
\frac{dV}{dt_r} = N \frac{dV_1}{dt_r} + (1 - N) \frac{dV_2}{dt_r} + \lambda \frac{dT}{dt_r}
\]

\[
= -Nu'_r w_{res} + 0 + \lambda \left(1 - \tau^N \eta_r - \tau^F \epsilon_r\right) w_{res} N
\]

\[
= - \left[u'_r - (1 - \tau^N \eta_r - \tau^F \epsilon_r) \lambda\right] w_{res} N
\]

and

\[
\frac{dV}{db} = N \frac{dV_1}{db} + (1 - N) \frac{dV_2}{db} + \lambda \frac{dT}{db}
\]

\[
= Nu'_h (1 - e) s + 0 - \lambda \left(1 - \tau^N \eta_h + \tau^F \epsilon_h\right) (1 - e) s N
\]

\[
= \left[u'_h - (1 - \tau^N \eta_h + \tau^F \epsilon_h) \lambda\right] (1 - e) s N
\]

and

\[
\frac{dV}{ds} = N \frac{dV_1}{ds} + (1 - N) \frac{dV_2}{ds} + \lambda \frac{dT}{ds}
\]

\[
= -N (1 - t_1 - \nabla \sigma) u'_1 s + 0 + \lambda \left(1 - t_1 - \tau^S \sigma - \tau^N \eta_s\right) s N
\]

\[
= - \left[(1 - t_1 - \nabla \sigma) u'_1 - (1 - t_1 - \tau^S \sigma - \tau^N \eta_s) \lambda\right] s N
\]

and

\[
\frac{dV}{dm} = N \frac{dV_1}{dm} + (1 - N) \frac{dV_2}{dm} + \lambda \frac{dT}{dm}
\]

\[
= -N\phi' \xi s + 0 - \lambda \left(k - \tau^N \eta_m - \tau^F \epsilon_m\right) s N
\]

\[
= \left[-\phi' \xi - (k - \tau^N \eta_m - \tau^F \epsilon_m) \lambda\right] s N
\]

as well as

\[
\frac{dV}{dt_2} = N \frac{dV_1}{dt_2} + (1 - N) \frac{dV_2}{dt_2} + \lambda \frac{dT}{dt_2}
\]

\[
= 0 - (1 - N) u'_2 w_2 + \lambda \left(1 + \tau^N \eta_2\right) w_2 (1 - N)
\]

\[
= - \left[u'_2 - (1 + \tau^N \eta_2) \lambda\right] w_2 (1 - N)
\]

QED.
Proof (proposition 1): we provide proof for parts (a) to (e) of the proposition in turn. For each part, we use optimality conditions as stated in lemma 1.

• Part (a): there is limited unemployment insurance:

\[
\frac{u'_r}{u'_h} = \frac{1 - \tau^E \varepsilon_r}{1 + \tau^E \varepsilon_b} < 1 \quad u_r > u_h
\]

Dividing the second condition of the lemma \(\lambda = u'_r/(1 - \tau^N \eta_r - \tau^E \varepsilon_r)\) by the third condition \(\lambda = u'_h/(1 - \tau^N \eta_h + \tau^E \varepsilon_b)\) delivers

\[
\frac{u'_r}{u'_h} = \frac{1 - \tau^N \eta_r - \tau^E \varepsilon_r}{1 - \tau^N \eta_h + \tau^E \varepsilon_b}
\]

Using \(\eta_r = u'_r/N_i'\) and \(\eta_h = u'_h/N_i'\), one has

\[
\frac{u'_r}{u'_h} = \frac{1 - \tau^N u'_r/N_i' - \tau^E \varepsilon_r}{1 - \tau^N u'_h/N_i' + \tau^E \varepsilon_b}
\]

so

\[
u'_r (1 - \tau^N u'_h/N_i' + \tau^E \varepsilon_b) = u'_h (1 - \tau^N u'_r/N_i' - \tau^E \varepsilon_r)\]

and

\[
u'_r (1 + \tau^E \varepsilon_b) = u'_h (1 - \tau^E \varepsilon_r)\]

Thus, using the fact that elasticities \(\varepsilon_j > 0\) and \(\tau^E = t_r w_r + b > 0\),

\[
\frac{u'_r}{u'_h} = \frac{1 - \tau^E \varepsilon_r}{1 + \tau^E \varepsilon_b} < 1.
\]

The concavity of the utility function concludes.

• Part (b) there is overproportional earnings risk insurance:

\[
\frac{u'_r}{u'_1} = 1 - \tau^E \varepsilon_r < 1 \quad u_r > u_1
\]

As for part (a) but using the first condition of the lemma \(\lambda = u'_1/(1 - \tau^N \eta_1)\) instead of the third condition and using \(\eta_1 = u'_1/N_i'\).

• Part (c) optimal tax redistribution is limited: when \(\tau^N > 0\),

\[
\frac{u'_1}{u'_2} = \frac{1 - \tau^N \eta_1}{1 - N} < 1 \quad u_1 > u_2
\]
A. PROOF DETAILS

Dividing the first condition of the lemma \( \lambda = u' / (1 - \tau^N \eta_1) \) by the last condition \( \lambda = u' / (1 + \tau^N \eta_2) \) delivers

\[
\frac{u'}{u''} = \frac{1 - \tau^N \eta_1}{1 + \tau^N \eta_2}.
\]

Using \( \eta_1 = u' / Ni' \) and \( \eta_2 = u' / (1-N)i' \) gives

\[
\frac{u'}{u''} = \frac{1 - \tau^N u' / Ni'}{1 + \tau^N u' / (1-N)i'}
\]

so

\[
u' (1 + \tau^N u'' / (1-N)i') = u'' (1 - \tau^N u' / Ni')
\]

\[
= u'' (1 - \tau^N u' (1-N) / (1-N)Ni')
\]

\[
= u'' (1 - \tau^N u' / (1-N)Ni' + \tau^N u' / (1-N)i')
\]

and

\[
u' = u'' (1 - \tau^N u' / (1-N)Ni') = u'' (1 - \tau^N \eta_1 / (1-N))
\]

thus, using the assumption that \( \tau^N > 0 \) and noting that \( \eta_1 = u' / Ni' > 0 \),

\[
\frac{u'}{u''} = 1 - \tau^N \eta_1 / (1-N) < 1.
\]

The concavity of the utility function concludes.

\* Part (d) employment protection internalizes firing cost externalities:

\[
t_s = t_1 w_1 + [(1-e) b - t_r w_r e] + km + \nabla
\]

Using \( \eta_s = (1 - t_1 - \nabla \sigma) \eta_1 \), the first condition of the lemma \( \lambda = u' / (1 - \tau^N \eta_1) \) and the fourth condition \( \lambda = (1 - t_1 - \nabla \sigma) u' / (1 - t_1 - \tau^S \sigma - \tau^N \eta_s) \), we have

\[
u' / (1 - \tau^N \eta_1) = u' \eta_s / (1 - t_1 - \tau^S \sigma - \tau^N \eta_s) \eta_1
\]

so

\[
(1 - \tau^N \eta_1) \eta_s = (1 - t_1 - \tau^S \sigma - \tau^N \eta_s) \eta_1
\]

and

\[
\eta_s = (1 - t_1 - \tau^S \sigma) \eta_1
\]

that is

\[
(1 - t_1 - \nabla \sigma) \eta_1 = (1 - t_1 - \tau^S \sigma) \eta_1
\]

or

\[
\nabla = \tau^S.
\]
Using the definition \( \tau^S = t_s + [e_t w_r - (1 - e) b] - km - t_1 w_1 \), one has
\[
\nabla = t_s + [e_t w_r - (1 - e) b] - km - t_1 w_1
\]
so
\[
ts_s = t_1 w_1 + [(1 - e) b - e_t w_r] + km + \nabla.
\]

- Part (e) implementation of active labor market policies is optimal:
\[
\frac{du^e}{dm} = \frac{-\phi' \zeta}{u_1'} = k - \tau^E \epsilon_m
\]
Dividing the fifth condition of the lemma \( \lambda = -\phi' \zeta / (k - \tau^N \eta_m - \tau^E \epsilon_m) \) by the first condition \( \lambda = u'_1 / (1 - \tau^N \eta_1) \), one has
\[
\frac{-\phi' \zeta}{u'_1} = \frac{k - \tau^N \eta_m - \tau^E \epsilon_m}{1 - \tau^N \eta_1}.
\]
Using \( \eta_1 = u'_1 / N_i' \) and \( \eta_m = -\phi' \zeta / N_i' \) gives
\[
-\phi' \zeta \left(1 - \tau^N u'_1 / N_i' \right) = u'_1 \left(k + \tau^N \phi' \zeta / N_i' - \tau^E \epsilon_m \right)
\]
that is
\[
-\phi' \zeta = u'_1 \left(k - \tau^E \epsilon_m \right)
\]
or
\[
\frac{-\phi' \zeta}{u'_1} = k - \tau^E \epsilon_m.
\]
Using the envelope theorem on \( u^e = \max_e e \cdot u ((1 - t_r) w_r) + (1 - e) \cdot [u (h + b) - \chi] - \phi (m) \zeta (e) \) concludes.
QED.

### B Calibration details

The Appendix informs about calibration. Since job separation is concentrated in the volatile sector, the firing rate is much higher than the average value of 10% reported in Blanchard and Portugal (2001), for example. We calibrate the entry cost \( f_e \), the productivity parameter \( A \) and the dispersion \( z \) of the uniform distribution such that the equilibrium supports the separation rate, the wage structure and the zero profit condition. We normalize the wage structure by setting \( w_2 = 1 \). Workers in sector 1 earn the most and, if reemployed and retrained, earn more than workers who avoid the sector specific training and directly go to the safe sector. We calibrate the model to yield \( w_1 = 1.4 \) in the baseline scenario and set \( w_r = 1.15 \). The wage loss after job separation is roughly
B. CALIBRATION DETAILS

in line with the results noted in Burda and Mertens (2001), among others. They find that workers in the highest earnings quartile loose on average 17% of the last wage in their next job. We use this result to reflect our assumption that the volatile sector is more innovative and skill intensive.\footnote{Data from the ILO (2010) show that computer programmers in private insurance companies were paid 20% more than those in public administration (in 2006; US). Burda and Mertens (2001) also report that, on average across all earning classes, workers loose only 3.6% of the last wage in the next job in Germany, compared to a loss of 10-25% in the U.S.}

We are not aware of any empirical estimates of the value of home production and set \( h = 0.2 \times w_r \) together with \( b = 0.5 \times w_r \). While the replacement rate in UI benefits is roughly OECD average, the total replacement rate including the value of home production is \( 0.7 \) and, thus, relatively high. We set the firing tax to zero in the initial equilibrium while Table 1 reports optimal firing taxes as a share of annual salary \( (t_s/w_1) \) around 71-77%. These optimal values are somewhat higher than actual job dismissal costs equal to 6 months salary (midway from values in Hopenhayn and Rogerson, 1993, and Heckman and Pages, 2000). Surveys on the effect of ALMP find they are small and cost around one percent of GDP. The optimal values of ALMP spending fluctuate around \( 0.7\% \) of GDP in Table 1, ranging from \( 0.2 \) to \( 1.1\% \) depending on the specific scenario. With respect to the three behavioral margins, we first note that the firing elasticity \( ds = -\sigma s dt_s \) given in equation (5) is \( 0.2 \), meaning that an increase in the firing tax by 10 percentage points ceteris paribus reduces the separation rate by 2 percentage points.

The elasticity of job search is governed by the curvature of the effort cost function, but also by the utility loss of unemployment. Blanchard and Tirole (2008) mention evidence of this being substantial, so we set it at \( \chi = 0.3 \times w_r \). Given this value, we calibrate the concavity of the effort cost function \( \zeta \) to reflect empirical studies on the elasticity of unemployment spells with respect to UI benefits (0.5 in the survey by Krueger and Meyer, 2002). The calibration implies that, ceteris paribus, and starting from a value of \( b_r = 0.5 \), an increase in the replacement rate by 1 percentage point reduces the unemployment rate by \( 0.5 \) percentage points. The entry margin can either be interpreted as job creation or as training for sector specific skills. We parametrize the cost function \( i(N) \) such that a one percent increase in expected utility of entering sector 1 attracts additional workers of \( \eta \) percentage points of the labor force, \( dN = \eta \cdot dV_1/|V_1| \). We set the semi-elasticity at \( \eta = 0.8 \). Finally, results on optimal unemployment insurance are fundamentally driven by the degree of relative risk-aversion. Chetty (2008) uses a value of 1.75 and reports that values up to 5 are not implausible. We settle for a baseline value equal to \( \rho = 1.75 \).
Bibliography


[38] Kluve, Jochen (2006), The Effectiveness of European Active Labor Market Policy, *RWI Discussion Paper 37, RWI Essen*


Chapter 2

Neoclassical labor supply, heterogeneous agents and changing skill premiums

Thomas Davoine\(^1\)

New chapter, May 2013 (Updates August 2013)

Abstract

This paper analyzes cross-sectional properties of labor supply in a dynamic neoclassical model with several household classes separated by an exogenous and time-varying skill difference. Empirical evidence is mixed: high-wage earners worked less than low-wage earners in the US early in the 20th century and more after 1970; high-wage earners work more in some countries and the reverse in other countries; the latter case appears to be a contradiction with small but positive individual labor supply elasticities with respect to net wages. I jointly analyze growth and labor supply to exhibit a relative intertemporal labor supply effect: high-skilled workers have an incentive to increase labor supply relative to low-skilled workers when the skill premium increases. I show that over the long run, high-skilled workers provide more labor than low-skilled workers if the skill premium increases long enough, the same amount of labor for a constant premium and less labor if the premium decreases long enough. Under partial equilibrium, the theory also predicts that labor supply is reduced shortly after labor income tax increases, consistent with micro-evidence on individual net wage elasticities.

JEL classification: E13, E2, H31, J2

Keywords: neoclassical theory, intertemporal labor supply, heterogeneous agents, skill premium, labor income tax

\(^1\)Institute for Advanced Studies (IHS), Vienna, Austria and University of St.Gallen, Switzerland. Email: thomas.davoine@ihs.ac.at. Email: thomas.davoine@unisg.ch.
1 Introduction

Empirical evidence on the cross-section dimension of labor supply is mixed: the slope of the labor supply curve is sometimes positive and sometimes negative. In some countries, high-wage earners work more while low-wage earners work more in other countries. Patterns have slowly changed in the United States over the 20th century: low-wage earners worked more early in the century but less after 1970. There is no elementary extension of the neoclassical labor supply model to explain these empirical patterns. The goal of this paper is to provide such an extension. I consider households who differ in skill for exogenous reasons and the impact of skill premium variations on relative labor supply. I jointly analyze growth and labor supply and show that over the long run and in general equilibrium, high-skilled workers provide more labor than low-skilled workers if the skill premium increases long enough, the same amount of labor for a constant premium and less labor if the premium decreases long enough. Theoretical predictions are consistent with stylized evidence.

I consider a simple extension of a neoclassical model to be able to analyze the cross-section dimension of labor supply. Households are identical except that they are born with different skill levels, which they keep forever. The skill premium creates a wage premium. There is a relative intertemporal labor supply effect when the skill premium exogenously varies over time. For instance, if high-skilled households know the premium will increase over time, they have an incentive to work less today and more tomorrow, relative to low-skilled households. If the skill premium variation takes place over a long enough period, high-skilled will end up working more than low-skilled and the labor supply curve will have a positive slope. Conversely, if the premium decreases long enough, the curve will have a negative slope. The latter outcome is not a contradiction with empirical evidence of a positive elasticity of individual labor supply with respect to net wages. Indeed, a tax increase may lead high- and low-skilled households to work less, but it does not prevent high-skilled to work less than low-skilled before and after the tax reform.

Figure 1 shows that US high-wage earners worked less than low-wage earners in the first part of the century, and more towards the end of the 20th century. Cross-section data analysis from the special issue of the Review of Economic Dynamics (2010) finds a negative correlation between hours and wages in the last 20 years of the century in most countries (Germany, Italy, Japan, Mexico, Russia, Spain and Sweden) but not in all countries (UK and US).

Heathcote, Storesletten, and Violante (2010) provide a theoretical analysis which can account for cross-section variations of labor supply in the US between 1970 and 2000. Adding exogenous variations in the skill premium, productivity shocks, incomplete markets, gender and endogenous education decisions to a neoclassical model, their quantitative study is able to match theory and data in several dimensions of the labor market and consumption. While labor supply cross-section variations are matched, averages
however differ. It is also difficult to identify the driver for the change.

Predictions from more basic theories are not clear. In its static version, the neoclassical textbook model is ambiguous about the sign of the labor supply curve. At low wage levels, the substitution effect dominates the income effect so the slope is positive. At high wage levels, the opposite takes place. There are no sign results from the dynamic version of the textbook theory, either: King, Plosser, and Rebelo (1988) provide conditions on preferences to have balanced growth with constant labor supply but there is no heterogeneity in wages.

The small and growing literature which considers skill and wealth heterogeneity in the dynamic neoclassical model provides some answers. Turnovsky and García-Peñalosa (2008) consider differences in initial capital and analytically show that asset-rich households work less\(^2\). Maliar and Maliar (2001) look at differences in initial capital and permanent skills and numerically find that skill-rich work less under standard preference parameters. In the same setting, García-Peñalosa and Turnovsky (2012) show that asset-rich households push their labor supply down and skill-rich households push it up, opening the door to income and wealth mobility. The slope of the labor supply curve depends on initial asset and skill distributions but the sign of the slope remains constant over time.

My analysis is an extension of García-Peñalosa and Turnovsky (2012). Restricting to identical initial asset and skill distributions, I add technological progress and time variation in the skill differences to characterize the slope of the labor supply curve in terms of skill premium changes. I show analytically that high-skilled households provide more

\(^2\)Turnovsky and García-Peñalosa (2008) analytically show that differences in initial assets are sufficient to explain why asset-rich households work less, and that borrowing constraints and incomplete markets, as in Pijoan-Mas (2006) and Marcet, Obiols-Homs, and Weil (2007), are not needed to have the results. The question I ask is similar: are productivity shocks, incomplete markets, endogenous education decisions and gender premiums needed to have an increasing correlation between wages and hours?
labor when the skill premium increases long enough over time, and vice-versa. Skill premium variations qualitatively explain why US high-skilled worked less than low-skilled in the first part of the 20th century and more in the second part.

As Goldin (1999b) documents, the major transformations in the 20th century US education system were a strong development of secondary education before WWII and a moderate development of tertiary education afterward. The skill premium gap is thus likely to have decreased before WWII and increased afterward. The relative intertemporal labor supply effect from the theory predicts that low-skilled households increased their labor supply over time, as the skill premium decreased, ending up working more than high-skilled in the first part of the century. Conversely, high-skilled ended up working more in the second part of the century, when the premium increased.

I also consider labor supply variations after unexpected labor income tax increases and show that positive empirical elasticities of individual labor supply are not a contradiction with the negative slope of the labor supply in its cross-section dimension, when the skill premium decreases. A difference must be made between the cross-section properties of labor supply curves and individual responses to net wage variations.

The paper is related to other strands of the literature which consider further extensions to the basic neoclassical framework and intertemporal labor supply decisions, such as inflation, home production, institutions, incomplete insurance markets and imperfect labor markets. The next section provides more details. Section 3 describes the model, followed by the theoretical analysis section. Section 5 provides a numerical illustration while section 6 compares theory and stylized cross-section evidence in the US. I conclude in section 7.

2 Literature overview

In this section I provide an overview of the literature on intertemporal labor supply decisions and relate it to the goal of this paper, comparing labor supply of agents with different skill levels. I start with the textbook neoclassical model and then consider additions of wealth and skill heterogeneity, inflation, home production, institutions, incomplete insurance markets and imperfect labor markets. I finish with empirical evidence. For the sake of completeness, some of the information contained in the introduction will be repeated.

Textbook neoclassical theory: the real business-cycle literature introduced labor supply decisions into the neoclassical framework, assuming disutility of labor. King, Plosser, and Rebelo (1988) characterized preferences leading to balanced growth with constant labor supply. There is however no heterogeneity among households. This paper provides an extension of the balanced growth result when households differ in skills.
Wealth and skill heterogeneity: Turnovsky and García-Peñalosa (2008) analyze the effect of initial wealth differences in a neoclassical model with endogenous labor supply and show, as Sorger (2000), that wealthier households work less. As a consequence, wealth inequality is reduced over time. The result is driven by the declining marginal utility of wealth. As they have a lower marginal utility, wealthy choose to consume more of all goods, including leisure. In support of this result, authors mention "compelling empirical evidence" (p.1401). They also show that average labor supply declines over time in expanding economies. Without any skill difference, the model has however no labor supply curve expressed as a function of the wage.

Maliar and Maliar (2001) add skill heterogeneity as in Kydland (1984). Households are born with an exogenously given and permanent skill level. They numerically find that high-skilled work less and analytically prove that standard preferences can not be consistent at the same time with two US facts taking place in recent decades, namely high-skilled working more and wage differentials being smaller or equal to wealth differentials. Considering long-run dynamics, the present paper analytically prove that high-skilled do not work more when the skill premium is constant.

García-Peñalosa and Turnovsky (2012) use the same setting and extend the analysis to long-run dynamics without technological progress. Ceteris paribus, higher steady-state wealth pushes labor supply down while higher permanent skill push it up. The main novelty is that differences in skills now can lead to mobility. With initial assets heterogeneity only, endogenous labor supply leads to changes in wealth and income inequality over time, but there is no catch up phenomenon – someone born rich will remain richer always. With skill differences, an asset-poor household with sufficiently higher skills earns more labor income and can end up wealthier, both in wealth and income terms. The main difference with the present paper is that I provide a characterization of labor supply by skill level (labor supply curve) with exogenous technological progress, consider the special case of identical differences between initial wealth and skill\(^3\), and finally allow for variations in the skill premium.

The model in this paper is similar to the models presented so far. Contributions of this paper thus belong to this strand of the literature. The remaining literature adds other components to the theory. One contribution of this paper is that it can obtain similar results as this richer literature in a simpler theoretical setting, helping to identify which parts of the theory play a prominent role. In the continuation, I review this literature with richer settings.

Inflation: Lucas and Rapping (1969) noted and dismissed an apparent contradiction which is related to the one mentioned in the introduction. Growth theory assumes that

\(^3\)For instance, their equations (16) and (17) conclude that labor supply is the same for all skill levels when relative initial assets (their \(k_{0,i}\)) is the same as relative skill difference (their \(a_i\)). My proposition 1 provides the same result, considering the special case of constant skill premium.
per capita labor supply is insensitive to changes in real wages while the business cycle literature assumes that there is a positive correlation. The first theory thus suggests no slope to the labor supply curve while the second one suggests a positive slope. They add non-time separable utility of consumption and leisure as well as adaptive expectations on inflation into an otherwise standard neoclassical model. Estimation of the model results in labor supply elasticities of 0.03 over the long-run and 1.40 over the short run, providing support for both assumptions in growth and business cycle theories and showing that there is no contradiction between them. In this paper, I abstract from inflation and reduced-form specifications.

**Home production:** Rios-Rull (1993) shows that costly education and home production alternatives can explain the secular rise in education and why high-skilled households provide more labor supply: as they have a higher opportunity cost for home production, they prefer to work more on the market and acquire more goods on the market than at home. Cobb-Douglas preferences imply that leisure and consumption goods are perfect substitutes. Vandenbroucke (2009) adds a separation between general goods and leisure goods, the later requiring leisure time to be consumed. Technological progress raises real wages and makes leisure goods more affordable, explaining the decline in per capita labor working hours in the US between 1900 and 1950. Households who choose more education have higher skills, a higher wage and can afford the consumption of more leisure goods and thus provide less labor. Unlike Rios-Rull (1993), there is some complementarity between leisure and consumption goods. In this paper, I abstract from home production. Analyses are complementary.

**Institutions:** the decline in per capita labor supply and why the decline differs between countries has been investigated in neoclassical models with taxes and representative agents. Prescott (2004) shows that tax increases can explain why labor per capita declines over time and why it did so more in Europe than in North America. Ohanian, Raffo, and Rogerson (2008) extend the analysis to more countries and subsistence consumption and obtain similar results. This paper is complementary and provide analytical results with heterogeneous agents.

**Incomplete insurance markets:** here is a small and growing number of papers from the literature on heterogeneous agents and incomplete markets which analyze the insurance properties of labor supply decisions, on top of precautionary savings (for a recent review, see Heathcote, Storesletten, and Violante, 2009). Flodén (2006) identifies the channel. Pijoan-Mas (2006) and Marcet, Obiols-Homs, and Weil (2007) quantify the

---

4Blanchard (2004) and Alesina, Glaeser, and Sacerdote (2005) remark that the elasticity of labor supply used by Prescott (2004) is large and consistent with macro but not micro evidence. Shimer (2009) argues for a middle ground and concludes that tax increases account for a large part of the evidence, even if not all of it.
wealth effect on labor supply, when there are idiosyncratic productivity shocks and borrowing constraints. Households subject to persistent negative shocks have low assets and increase work hours to help smooth consumption, as they have a high marginal utility of consumption and in spite of low returns to work. In the present paper, I also identify a wealth effect on labor supply, albeit transitory. The effect is due to policy changes rather than idiosyncratic productivity shocks.

Heathcote, Storesletten, and Violante (2010) account for several empirical evidence of the labor market and consumption between 1970 and 2000 in the US. Adding gender and endogenous education decisions, their quantitative analysis is in particular consistent with changes (but not averages) in the correlation between hours and wages, taken skill premium variations as input. In this paper, I provide analytical results and show that skill premium variations alone are qualitatively consistent with the observed changes in the hours-wage correlation.

**Imperfect labor markets:** there is a small but increasing number of studies with incomplete insurance markets and a richer structure on the labor market, in particular imperfect labor markets and search frictions, such as Krusell, Mukoyama, and Sahin (2010). Many of these studies focus on unemployment insurance or business cycle facts, while the present paper focus on heterogeneity in a standard neoclassical setting without uncertainty.

Michelacci and Pijoan-Mas (2012) for instance obtain comparable results to Pijoan-Mas (2006) with imperfect labor markets, a job search friction and uncertainty on job offers (rather than productivity): households with more assets work fewer hours. They also show that exogenous wage inequality, which has been increasing in the US over the recent decades, explain why average work hours per capita has increased: longer hours increase human capital and the probability of obtaining good jobs, which are becoming better and better. The theory can also explain why per capita labor supply has not risen in Europe, as wage inequality has not increased.

**Empirical evidence:** there is a vast empirical literature which analyzes labor supply. Empirical studies which take the intertemporal dimension of household decisions into account find positive elasticities of labor supply with respect to net wages. While the conventional wisdom is that elasticities are small (Blundell and MaCurdy, 1999), recent analysis adding human capital dimensions and separating intensive and extensive margins argue that they can be larger (Keane and Rogerson, 2012). In the present paper, the analysis does not depend on the size of the elasticity parameter in the disutility of labor function and is thus consistent with both views.

Empirical evidence which compares decisions between households, rather than decisions by the same household in different conditions, shows more variation over time.

---

5See Heathcote, Storesletten, and Violante (2009) for other references.
Cross-section data in Costa (2000) show that US low-wage earners worked more hours than high-wage earners in the 1890’s and, to a lower extent, in the 1970’s, but fewer hours in the 1990’s (see figure 1 for a glimpse). Gimenez-Nadal and Sevilla (2012) tentatively finds similar reversal patterns between the 1970’s and the 1990’s in seven other industrialized countries, comparing low-educated and high-educated individuals. Heathcote, Storesletten, and Violante (2010) document a rising correlation between males wages and hours in the US between 1970 and 2000, starting from negative values in 1970 and becoming positive by 1980, an empirical finding which extends that of Costa (2000) until the end of the 20th century. This finding is verified in a number of other developed countries (see Krueger, Perri, Pistaferri and Violante, 2010, and papers from the same special issue of the journal).

3 Model

I jointly analyze growth and labor supply in a Ramsey neoclassical growth theory model for a finite number of heterogeneous classes of agents, with endogenous labor supply decisions and disutility of labor. Preferences are separable in consumption and leisure and consistent with balanced growth paths. For ease of reading, I expose the model with two classes but results extend to $N \geq 2$ classes.

Population: there is a mass one of infinitely-lived households who belong to one of two classes, high or low skilled. Each household keeps its skill level forever. High-skilled earn a higher wage than low-skilled, so will be richer. I denote them by $R$ and denote low-skilled, or poor households, by $P$. The size of each class, $N_R$ and $N_P$, remains constant. As the focus of the analysis is on the consequences of inequality, I do not explain the origin of the skill differences between the two classes of agents.

Production and labor market: There is a representative firm and perfect competition, so factors earn their marginal product. Total capital $K$ is the sum of households assets, $N_R a_R + N_P a_P$, and depreciates at rate $\delta$. Technology $A$ is labor augmenting and grows at an exogenous rate $g$. High-skilled households are more productive, providing $AA_R l_R$ units of effective labor, while low-skilled households provide $AA_P l_P$ units of effective labor, with $A_R$ and $A_P$ such that $0 < A_P < A_R$. I assume perfect substitution between high-skilled effective labor and low-skilled effective labor. I denote $L = N_R A_R l_R + N_P A_P l_P$.

\footnote{Although there is no full consensus on the reasons for wage inequality increases in several industrial countries since the 1970’s, one generally accepted contributing factor is education (supply of skills; see for instance Goldin and Katz, 2009). To some extent, education is a force which is exogenous to the labor market but contributes to shape it. Assuming exogenous variations of skills and impact on wages is thus a reasonable starting point.}
and, for ease of reading, call this value the aggregate labor supply\textsuperscript{7}. Production equals 
\[ Y = F(K, AL), \] where \( F \) is homogenous of degree one and has a Cobb-Douglas form with capital-share \( v \). The net interest rate is 
\[ r = f'(k) - \delta \] where \( f(k) = f(K/AL) \equiv F(K/AL, 1) = k^v. \] 
For the benefit of some analytical results, I make the technical assumption that the initial capital-labor ratio is no less than unity (\( k_0 = K_0/(A_0L_0) \geq 1 \)), which holds by a wide margin in all developed countries.

It is convenient to apply a normalization and use notation such that the wage rate of high-skilled equals 
\[ w^R = (1 + \alpha^R)w \] and the wage rate of low-skilled equals 
\[ w^P = (1 - \alpha^P)w, \] for \( \alpha^R, \alpha^P > 0 \), where 
\[ w = \partial F(K, AL)/\partial L \] is the marginal product of labor or aggregate wage rate. Using the notation \( \alpha \), one has
\[
\begin{align*}
  w^R &= (1 + \alpha^R)w \quad \Leftrightarrow \quad w^R = (1 + \alpha)w^P \\
  w^P &= (1 - \alpha^P)w \quad \Leftrightarrow \quad 1 + \alpha \equiv \frac{1 + \alpha^R}{1 - \alpha^P} > 1
\end{align*}
\]
One can choose measurement units such that either \( \alpha^R \) or \( \alpha^P \) is close or equal to zero, but not both. For technical reasons, I choose \( \alpha^R \) to be close to zero. Details of the normalization are included in the appendix. The skill difference between the two classes of households thus leads to a wage premium \( \alpha > 0 \) for the high-skilled. Following Heathcote, Storesletten, and Violante (2010), Michelacci and Pijoan-Mas (2012) and others, I consider the premium as an exogenous input. Because I focus on the impact of skill differences alone, I assume that initial assets are distributed in the same manner as skills\textsuperscript{8}, \( a^R_0 = (1 + \alpha)a^P_0 \).

**Preferences:** Following specifications in labor market theory from Shimer (2009) and Prescott, Rogerson, and Wallenius (2009), preferences are separable in consumption and disutility of labor, namely 
\[ u_0(c_t, l_t) = u(c_t) - v(l_t) \] with \( v \) convex increasing\textsuperscript{9}. I consider isoelastic functional forms 
\[ u(c) = c^{1-\theta}/(1-\theta) \] and 
\[ v(l) = \lambda l^{1+\varepsilon}/(1+\varepsilon), \] with \( \theta, \lambda, \varepsilon > 0 \). When \( \theta = 1 \), \( u(c) = \ln(c) \). Households discount the future at rate \( \rho \).

King, Plosser, and Rebelo (1988) show that \( \theta \) must equal 1 to have balanced growth paths with constant labor supply, so I will assume\textsuperscript{10} that \( \theta = 1 \). The parameter \( \varepsilon \) is the inverse of the (Frisch) elasticity of labor supply: microeconomic estimates suggest a high value, between 2 and 10, while macroeconomic estimates hint at a low value, around 0.25 (see Shimer [31] for references and a discussion, arguing for an intermediate value of 1). The actual value will not matter for the analytical results.

\textsuperscript{7}Strictly speaking, \( L \) contains productivity components, \( A^R \) and \( A^P \), and its name should reflect it. I abstract from defining a new name, as the name itself plays no role in the analysis.

\textsuperscript{8}See García-Peñalosa and Turnovsky (2012) for an analysis of outcomes when the skill premium is constant but the initial asset distribution and skill premium distribution may differ.

\textsuperscript{9}One obtains the same results when considering preferences separable in consumption and utility of leisure, 
\[ u_0(c_t, l_t) = u(c_t) + v(1-l_t) \] with \( v \) concave increasing.

\textsuperscript{10}Some of the results from this paper hold with \( \theta \neq 1 \). I discuss these cases in the appendix.
CHAPTER 2. LABOR SUPPLY AND SKILL PREMIUMS

Adding a budget constraint and a no-Ponzi terminal condition, each class $i \in \{R, P\}$ of households solves the following maximization problem:

$$\max_{c^i(t), l^i(t)} \int_0^\infty \exp(-rt)u_0(c^i(t), l^i(t))dt$$

s.t.

$$\dot{a}^i(t) = r(t)a^i(t) + w^i(t)l^i(t) - c^i(t)$$

$$\lim_{t \to \infty} \left[ a^i(t) \exp \left( -\int_0^t r(s)ds \right) \right] \geq 0.$$ 

4 Analysis

I start with some intuitive building blocks. To allow for comparisions and to build intu-ition in a simple manner, I derive these building blocks from textbook theory.

Static textbook labor supply theory is built in partial equilibrium but applies to repre-sentative and heterogeneous agents. In this theory, the cross-section slope of the labor supply curve is ambiguous. For low wages, the marginal utility of consumption is high and the substitution effect dominates: higher wages provide a higher reward to labor, whose supply is larger. The slope is positive. At higher wages, the income effect dom-inates: higher wages provide more income, increasing the consumption of both goods, consumption and leisure, so the labor supply declines. The slope is negative. Figure 2 provides the resulting partial equilibrium and static theory version of the labor supply.

What to expect in a dynamic and general equilibrium context? When preferences are consistent with balanced growth, representative agents provide constant labor supply over the long run, even if real wages increase with technological progress (King, Plosser, and Rebelo, 1988). The substitution and income effect cancel out over the long run.

I will provide below an extension of this result to heterogeneous agents: when the skill premium is constant, high-skilled and low-skilled households also provide constant labor supply. Moreover, they provide the same labor. When the skill premium is not constant, relative intertemporal labor supply effects appear, which I will discuss below. High- and low-skilled households may not provide the same labor.
The section is divided in two parts. First I consider the cross-section dimension of the labor supply curve. Then I consider productivity shocks effects on the curve.

4.1 Cross-section of dynamic labor supply

The analysis starts with optimality conditions for each household class. For ease of reading, I only write the class index $i$ and time $t$ when needed to avoid confusion. The current-value Hamiltonian is $\mathcal{H}(c, l, a, \mu) = u_0(c, l) + \mu (ra + wl - c)$. Optimality conditions $\partial \mathcal{H} / \partial c = 0$, $\partial \mathcal{H} / \partial l = 0$ and $\partial \mathcal{H} / \partial a + \mu - \rho \mu = 0$ and differentiation lead to the following two relationships:

$$\frac{\partial u_0(c, l)}{\partial l} + \frac{\partial u_0(c, l)}{\partial c} w = 0$$

$$r - \rho = \theta \frac{\dot{c}}{c} - \left( \frac{\partial^2 u_0(c, l)}{\partial c \partial l} \right) w \frac{\dot{l}}{l}.$$

The second relationship is an adjusted Euler equation. When labor supply is constant ($\dot{l} = 0$), the relationship is the standard Euler equation. Using these conditions, I derive labor supply curves for each household class and then compare them. With $\theta = 1$, functional forms $u(c) = \ln(c)$ and $v(l) = \lambda l^{1+\varepsilon}/(1+\varepsilon)$ in preferences $u_0(c_t, l_t) = u(c_t) - v(l_t)$, we have $\partial u_0(c, l)/\partial l = -v'(l) = -\lambda l^{\varepsilon}$, $\partial u_0(c, l)/\partial c = u'(c) = c^{-1}$ and $\partial^2 u_0(c, l)/\partial c \partial l = 0$. The first optimality condition is then

$$-\lambda l^{\varepsilon} + c^{-1} w = 0 \iff \lambda c l^{\varepsilon} = w \iff \left( \frac{1}{\lambda} \frac{w}{c} \right)^{\frac{1}{\varepsilon}} = l.$$

For each household class with a wage schedule $w_t$, the labor supply curve is thus

$$c_t = \frac{1}{\lambda} \frac{1}{l_t^{\varepsilon}} w_t \iff l_t = \left( \frac{1}{\lambda} \frac{w_t}{c_t} \right)^{\frac{1}{\varepsilon}}.$$

Assume now two classes of households, low and high skills, earnings wages $w_t^R = (1 + \alpha) w_t^P$ for a positive $\alpha > 0$ and consider the optimal labor supply schedules $l^R$ and $l^P$. Comparing them, we have

$$\frac{c^R}{c^P} = \left( \frac{1}{\lambda} \frac{1}{(l_t^R)^{\varepsilon} w_t^R} \right) / \left( \frac{1}{\lambda} \frac{1}{(l_t^P)^{\varepsilon} w_t^P} \right) = \left( \frac{l_t^P}{l_t^R} \right)^{\varepsilon} \frac{w_t^R}{w_t^P}.$$

The second optimality condition, with $\partial^2 u_0(c, l)/\partial c \partial l = 0$, is $r - \rho = \dot{c}/c$. The consumption growth rate is thus identical for all household classes. From the previous relationship with $w_t^R = (1 + \alpha) w_t^P$, it follows that $(l_t^P/l_t^R)^{\varepsilon} (1 + \alpha)$ is constant over time. The following lemma sums up the derivations:
**Lemma 1. (Optimality conditions):** optimal labor supply and consumption decisions are characterized by the following:

\[ l^i = \left( \frac{1}{\lambda} \frac{w^i}{c^i} \right)^{\frac{1}{\varepsilon}} \quad \left( \frac{l^P}{l^R} \right)^{\varepsilon} (1 + \alpha) = \text{constant} \]

\[ \frac{c^R}{c^P} = \left( \frac{l^P}{l^R} \right)^{\varepsilon} \frac{w^R}{w^P} \]

\[ \frac{\dot{c}^R}{\dot{c}^P} = \frac{\dot{c}^P}{\dot{c}^P} = (r - \rho) \]

These relationships are not new\(^{11}\) but key for the analysis. Consider the intermediate elasticity value \( \varepsilon = 1 \). Then

\[ l^i = \frac{1}{\lambda} \frac{w^i}{c^i} \quad \left( \frac{l^P}{l^R} \right)^{1 + \alpha} = \text{constant} \]

Ceteris paribus, higher wage \( w^i \) suggest a larger labor supply \( l^i \). Ceteris paribus analysis however can be misleading: a higher wage \( w^i \) should lead to higher consumption expenditures \( c^i \), reducing the likelihood of a large \( l^i \). Being the main novelty in this paper, it is interesting to consider skill premium variations. Assume that the premium \( \alpha \) increases over time. The second relationship says that labor supply \( l^P \) of the poor will increase slower than labor supply \( l^R \) of the rich, a relative intertemporal labor supply effect. Intuitively, it makes more sense for rich households to work more than poor households in the future rather than in the present, when they earn an even larger wage. If the premium increases for a long enough period, rich will eventually work more. The converse holds. This relative intertemporal labor supply effect is the main driver of the results provided below.

In continuation, I make a formal derivation of this intuitive discussion.

**Lemma 2. (Labor supply in the cross-section, non-increasing skill premium):** consider an economy with two agents classes separated by an exogenous and non-increasing skill premium differential (\( \dot{\alpha} \leq 0 \)). Then over the long run in any equilibrium, high-skilled households work less or as much as low-skilled households:

\[ l^R \leq l^P. \]

With representative agents, King, Plosser, and Rebelo (1988) have shown that there is a unique equilibrium. When there are several classes of households, this unique equilibrium outcome may or may not extend. Different initial assets \( a^i_0 \) for instance may lead to different equilibrium schedules \( l^i(t) \). Lemma 2 however holds even if there are

\(^{11}\)The condition (5.8) in Kydland (1984) is similar to the left bottom relationship of the lemma.
multiple equilibrium. Its proof exploits some properties of optimal solutions, without the need of a full characterization of the long run equilibrium. As a consequence of proposition 2 however it will turn out that the equilibrium is unique and independent from the transition path when the skill premium is constant

Note also that it is easy to extend the result to \( N \geq 2 \) classes of households; that the lemma also holds in a partial equilibrium setting where wage and interest rates are exogenously given, constant or not, in which case the lemma holds not only on the long run but at any point in time; and that the lemma hinges on the assumption of identical labor disutility parameter \( \epsilon \) for rich and poor households, which may not be true in reality.

I provide a sketch of the proof for the easier case of constant skill premium and refer to the appendix for details and the general case.

**Proof (sketch):** the idea is to assume \( \ell^R > \ell^P \) in the optimal allocations \((c^R_*, l^R_*, c^P_*, l^P_*)\) and exhibit a contradiction. I build the alternative allocation \((c^R_2, l^R_2, c^P_2, l^P_2)\) and show first that it is a strict Pareto improvement over \((c^R_*, l^R_*, c^P_*, l^P_*)\) and second that it is feasible for the (partial equilibrium) price schedule \((w^R_2, w^P_2, r_2) = (w^R_*, w^P_*, r_*)\). The complete proof then derives a related allocation which is also a Pareto improvement and feasible under general equilibrium.

Clearly the allocation leaves the poor indifferent. By lemma 1 and the contradiction hypothesis \( \ell^R > \ell^P \), we have

\[
\frac{c^R_*}{c^P_*} = \left[ \left( \frac{l^P_*}{l^R_*} \right)^\epsilon \frac{w^R_*}{w^P_*} \right] = \left[ \left( \frac{l^P_*}{l^R_*} \right) \right] (1 + \alpha) < (1 + \alpha),
\]

so \( c^R_* < (1 + \alpha) c^P_* = c^R_2 \). It follows that the second allocation \((c^R_2, l^R_2, c^P_2, l^P_2)\) is a strict Pareto improvement, as both the consumption utility increase and the labor disutility component decrease for the rich.

For poor households, neither the allocation nor the price schedule change so \((c^P_2, l^P_2) = (c^P_*, l^P_*)\) remains a feasible allocation. Keeping in mind the household maximization problem, the allocation \((c^i, l^i)\) is feasible for the schedule \((w_*, r_*)\) if and only if

\[
\dot{a}^i(t) = r_*(t) a^i(t) + w_*(t) l^i(t) - c^i(t)
\]

\[
\lim_{t \to \infty} \left[ \dot{a}^i(t) \exp \left( - \int_0^t r_*(s) ds \right) \right] \geq 0,
\]

where \( w^R_*(t) = (1 + \alpha R) w_* (t) \) and \( w^P_*(t) = (1 - \alpha P) w_* (t) \). We know that these conditions hold for \( i = P \) with \((c^P_*, l^P_*) = (c^P_2, l^P_2) = (c^P_2, l^P_2)\) and we need to verify they hold for \( i = R \) and \((c^R_*, l^R_*) = (c^R_2, l^R_2)\). Multiplying the first and second conditions for \( i = P \) by \((1 + \alpha) = (1 + \alpha R)/(1 - \alpha P)\); defining \( a^R(t) \equiv (1 + \alpha) a^P(t) \), consistent with the assumption that

\[\text{Homotheticity in consumption decisions, a consequence of lemma 1, is insufficient to guarantee the unicity of equilibrium in other cases, because some differences in labor supply decisions are left open by this analysis.}\]
CHAPTER 2. LABOR SUPPLY AND SKILL PREMIUMS

initial assets satisfy \( a^R_0 \geq (1 + \alpha)a^P_0 \); and reminding that \((1 + \alpha)c^P_i(t) = c^R_i(t)\) and \(i^R_i(t) = l^R_i(t)\), we have that the two conditions hold for \(i = R\) and the allocation \((c^R, l^R) = (c^R_2, l^R_2)\). The \((c^R_2, l^R_2, c^P_2, l^P_2)\) allocation is thus feasible.

QED.

Combining the previous results lead to the first unambiguous characterization of the slope of the labor supply curve, when the skill premium does not increase. The optimality condition \(\partial (l^P/l^R) \cdot (1 + \alpha) / \partial t = 0\) from lemma 1 says that labor supply of the high-skilled \(l^R\) grows at the same rate as labor supply of the low-skilled \(l^P\) when the skill premium \(\alpha\) is constant and that \(l^R\) grows slower than \(l^P\) when \(\alpha\) declines. Consider the case of a non-increasing skill premium \(\dot{\alpha} \leq 0\) which at least once is declining, \(\dot{\alpha}(t_0) < 0\). Since labor supply of the high-skilled \(l^R\) is never greater than labor supply \(l^P\) of the low-skilled (lemma 2), since \(l^R\) never grows faster than \(l^P\) in this case (lemma 1) at all points in time and since there is at least one point in time \((t_0)\) when \(l^R\) grows slower than \(l^P\), labor supply of the high-skilled \(l^R\) is strictly lower than labor supply of the low-skilled \(l^P\) for all subsequent points in time. Summing up, we have\(^{13}\)

**Lemma 3. (Labor supply in the cross-section, strictly non-increasing skill premium):** consider an economy with a skill premium \(\alpha\) which is either constant or declines at all points in time, \(\dot{\alpha} \leq 0\), and which declines at least once, \(\dot{\alpha}(t_0) < 0\). Then over the long run in any equilibrium, high-skilled households work strictly less than low-skilled households:

\[
\begin{align*}
    l^R(t) &\leq l^P(t) \quad \forall t \leq t_0 \\
    l^R(t) &< l^P(t) \quad \forall t > t_0
\end{align*}
\]

One can obtain symmetric results when the skill premium \(\alpha\) is non-decreasing, with some adjustments on the characteristics of the economy. Arguments and reasoning are similar to the non-increasing skill premium case.

**Lemma 4. (Labor supply in the cross-section, non-decreasing skill premium):** consider an economy with two agents classes separated by an exogenous and non-decreasing skill premium differential \((\dot{\alpha} \geq 0)\). Then over the long run in any equilibrium, high-skilled households work more or as much as low-skilled households:

\[l^R \geq l^P.\]

\(^{13}\)Formally, a consequence of lemma 1 is that \(l^R_i(t) \leq l^P_i(t)\) for all time \(t\) and \(l^R_i(t_0) < l^P_i(t_0)\). By lemma 2, \(l^R_i(t) \leq l^P_i(t)\) for all time \(t\). It is thus impossible to have \(l^R_i(t_0 + \Delta t) = l^P_i(t_0 + \Delta t)\) for small \(\Delta t > 0\), so \(l^R_i(t_0 + \Delta t) < l^P_i(t_0 + \Delta t)\). Because \(l^R_i/t^R_i(t) \leq l^P_i/t^P_i(t)\) also for all time \(t \geq t_0 + \Delta t\), we have \(l^R_i(t) < l^P_i(t)\) for all time \(t > t_0\).
The appendix provides the proof of the lemma, which is very similar to the proof for lemma 2. The next result follows from the same steps as lemma 3, with reversed signs:

**Lemma 5. (Labor supply in the cross-section, strictly non-decreasing skill premium):** consider an economy with a skill premium $\alpha$ which is either constant or increases at all points in time, $\dot{\alpha} \geq 0$, and which increases at least once, $\dot{\alpha}(t_0) > 0$. Then over the long run in any equilibrium, high-skilled households work strictly more than low-skilled households:

\[ l^R(t) \geq l^P(t) \quad \forall t \leq t_0 \quad l^R(t) > l^P(t) \quad \forall t > t_0 \]

Putting lemmas 2 and 4 together, we obtain an unambiguous cross-section characterization of the labor supply curve in a dynamic setting:

**Proposition 1. (Labor supply in the cross-section, constant skill premium):** consider an economy with two agents class with an exogenous and constant skill premium differential ($\dot{\alpha} = 0$). Then over the long run in any equilibrium, high-skilled households and low-skilled households provide the same labor supply:

\[ l^R = l^P. \]

These results characterize the cross-section dimension of dynamic labor supply in neoclassical theory. While the sign of the slope of the labor supply is ambiguous in static textbook theory (see introduction of this section), there is no ambiguity in lemmas 3, 5 and proposition 1: the slope is positive when the skill premium increases, null when the premium is constant and negative when the premium decreases.

The intuition behind these results is the following. Consider the case of initial constant skill premium ($\dot{\alpha} = 0$ for $t \leq t_0$) and subsequent increase ($\dot{\alpha} > 0$ for $t > t_0$), as in lemma 5. The lemma predicts that high-skilled and low-skilled households will initially provide the same labor supply ($l^R(t) = l^P(t)$ for $t \leq t_0$) and that high-skilled will then provide more labor ($l^R(t) > l^P(t)$ for $t > t_0$). These decisions come from their intertemporal nature. Before the increase in the premium, high-skilled know that they will obtain an even higher wage in the future, relative to the low-skilled. From an intertemporal point of view, it is thus in the future, after the increase in the premium, that it is worth working more than the low-skilled. The same intuition applies when the skill premium is decreasing, but reversed.

Figure 3 presents a graphical representation of dynamic labor supply decisions at a given time $t$, depending on skill premium variations. Just as figure 2 for the static case, it
provides work hours as a function of wages. It is based on lemmas 3, 5 and proposition 1 and illustrates the relative intertemporal labor supply effect of skill premium variations.

\[
\begin{align*}
\dot{\alpha} &> 0 \\
\dot{\alpha} &= 0 \\
\dot{\alpha} &< 0
\end{align*}
\]

Figure 3: Labor supply decisions depending on skill premium variations (time \(t\) snapshot)

One could wonder why there is no wealth effect on labor supply as in Turnovsky and Garcia-Peñalosa (2008) and Pijoan-Mas (2006) when the skill premium is constant: high-skilled households could consume their higher assets over both goods, consumption and leisure, reducing their labor supply. There is no such wealth effect for the same reasons that the income effect cancels out with the substitution effect in the dynamic textbook neoclassical model. Even if real wages increase with growth, preferences on consumption goods and leisure are such that the balanced growth path has constant labor supply, balancing income and substitution effects.

The dynamic textbook neoclassical theory result however does not extend to the case of changing skill premiums. Lemma 5 shows that high-skilled labor supply will increase faster than low-skilled supply as long as the skill premium is strictly increasing and that high-skilled will provide more labor.

In this case, high-skilled benefit not only from technological progress, as in the textbook theory, but also from the presence of the low-skilled. There is a form of redistribution from the low- to the high-skilled. The skill premium increase (\(\dot{\alpha} > 0\)) means that the labor effectiveness of the high-skilled is not only higher (\(A^{R}/A^{P}\)) but keeps increasing (\(\dot{A}^{R}/A^{R} > \dot{A}^{P}/A^{P}\)), relative to low-skilled. Ceteris paribus, the aggregate capital-efficient labor ratio is higher when the high-skilled belong to an heterogeneous economy than in an economy where they are alone. When the skill premium is constant (\(\dot{\alpha} = 0, \dot{A}^{R}/A^{R} = \dot{A}^{P}/A^{P}\)) there is only a level effect. When the premium increases (\(\dot{\alpha} > 0, \dot{A}^{R}/A^{R} > \dot{A}^{P}/A^{P}\)), there is also a growth effect. The aggregate capital-efficient labor ratio and high-skilled real wages are thus increasing faster in an heterogeneous economy than in an economy solely made of high-skilled. Income and consumption opportunities keep getting better and better for the high-skilled, who provide more and more labor. The opposite takes place for the low-skilled. Preferences parameters, which balance substitution effect and income effect in a representative economy, no longer balance these two effects in a heterogeneous agent economy with increasing skill differentials.

The previous results either assume a constant skill premium \(\alpha\) or monotonicity in the variations: either \(\dot{\alpha}(t) \leq 0\) for all \(t\) or \(\dot{\alpha}(t) \geq 0\). In reality, it is possible that the skill
premium increases over a stretch of time and decreases at other points in time. In this case, the main results obtained so far can not be applied. If we assume a degree of uncertainty on future changes of the skill premium and if we make a loose treatment of uncertainty, one can informally derive some testable implications:

Testable implications 1. (Labor supply in the cross-section, unpredictable skill premium variations): assume an economy with unpredictable and continuous variations in the skill premium \( \alpha \), such that households take recent increases (resp. decreases) in the skill premium as prediction of permanent future increases (resp. decreases). The theory predicts the following impacts of the skill premium variation on the cross-section characteristics of the labor supply:

(a) if the skill premium is constant and then increases (there is a \( t_0 \) such that \( \dot{\alpha}(t) = 0 \quad \forall t \leq t_0 \) and \( \dot{\alpha}(t) > 0 \quad \forall t > t_0 \)), initially high- and low-skilled households will work the same and then high-skilled work more

\[
l^R(t) = l^P(t) \quad \forall t \leq t_0 \quad \quad l^R(t) > l^P(t) \quad \forall t > t_0
\]

(b) if the skill premium is constant and then decreases (there is a \( t_0 \) such that \( \dot{\alpha}(t) = 0 \quad \forall t \leq t_0 \) and \( \dot{\alpha}(t) < 0 \quad \forall t > t_0 \)), initially high- and low-skilled households will work the same and then high-skilled work less

\[
l^R(t) = l^P(t) \quad \forall t \leq t_0 \quad \quad l^R(t) < l^P(t) \quad \forall t > t_0
\]

(c) if the skill premium is sometimes constant, sometimes increases and sometimes decreases, then high- and low-skilled households will not provide the same labor supply but high-skilled labor supply will increase faster when the skill premium increases, and vice-versa

\[
\frac{l^R(t)}{l^P(t)} > \frac{l^P(t)}{l^P(t)} \quad \text{when} \quad \dot{\alpha}(t) > 0 \quad \text{(and vice-versa)}
\]

(d) if the skill premium is sometimes constant, sometimes increases and sometimes decreases, but there is a prolonged period of skill premium increase, then high-skilled households are likely to provide more labor supply

\[l^R(t) > l^P(t) \quad \text{when} \quad \alpha > 0 \quad \text{long enough and after a while}\]

(e) if the skill premium is sometimes constant, sometimes increases and sometimes decreases, but there is a prolonged period of skill premium decrease, then high-skilled households are likely to provide less labor supply

\[l^R(t) < l^P(t) \quad \text{when} \quad \alpha < 0 \quad \text{long enough and after a while}\]

Testable implications (a) and (b) follow from proposition 1, lemmas 3 and 5: as long
as the premium is constant, households behave according to proposition 1, ignorant of future skill premium changes; after the premium unexpectedly changes (at time $t_0$), they behave according to either lemma 3 or lemma 5. The implication (c) is a consequence of lemma 1. Implications (d) and (e) are the only ones which are informally derived from the theory. From (c), we know that the labor supply of the high-skilled $l^R$ will increase more than that of the low-skilled $l^P$ when the skill premium increases. It could however be that it starts from a lower level so that a small period of skill premium increase is not sufficient for $l^R$ to catch up on $l^P$. When the increase in the skill premium takes place over a long enough period of time, then $l^R$ is likely to catch up and overcome $l^P$. For long enough periods of skill premium increase, case (d) is close to the case (a) and they lead to the same prediction.

In section 6, I will compare these testable implications with stylized evidence on education, the skill premium and the labor market. One may however already wonder if the theory is consistent with empirical evidence on the elasticity of labor supply with respect to net wages. The next subsection provides a discussion.

### 4.2 Exogenous wage shocks

Proposition 1 shows that in a cross-section and when the skill premium is constant, the labor supply curve has no slope. This may come as a surprise, as there is micro-evidence of a small but positive individual elasticity of labor supply with respect to net wages (see Blundell and MaCurdy, 1999, for a review). How can the two be consistent?

The following results help.

I consider unexpected shocks to net wages, such as labor productivity shocks or the introduction of labor income taxes. As variations of labor income taxes constitute a popular identification mechanism in empirical studies of the elasticity of labor supply, I present the analysis with labor income taxes. Results however also apply to productivity shocks. I assume that government finance its consumption with labor income taxes so that its budget is balanced in every period and that government consumption delivers no utility to households, following the public finance literature.

The result below shows a transitory wealth effect on labor supply under partial equilibrium, after a tax reform. Immediately after an unanticipated labor income tax increase, households have more assets than they would have built if they had been under the target tax regime all along. They use this extra wealth to purchase more consumption goods and more leisure, that is decrease labor supply. An econometric study comparing outcomes before and after the tax reform, an identification mechanism of the elasticity of labor supply with respect to wages, would then measure a positive elasticity.

The analysis shows that there is no contradiction between a positive individual labor supply elasticity and labor supply curves with negative or no slope in their cross-section dimension. Nothing prevents high-skilled to work less than low-skilled before and after
a tax increase, even if both types of households work less after the tax reform. The result holds with representative or heterogeneous agents with constant skill premiums. The heterogeneous agents result is built on the following proposition, which shows that households decisions are homothetic when the skill premium is constant:

Proposition 2. (Homotheticity and constant labor supply, constant skill premium): consider an economy with two agents class with an exogenous and constant skill premium differential (given by constant \( \alpha \)). Then households provide constant labor supply over the long run and high-skilled households are homothetic versions of low-skilled households:

\[
\begin{align*}
I^R &= I^P = \text{constant} \\
\frac{c^R}{c^P} &= \frac{y^R}{y^P} = \frac{a^R}{a^P} = (1 + \alpha) \equiv \frac{w^R}{w^P}
\end{align*}
\]

where \( y^i \equiv ra^i + w^i l^i \) denotes total income.

Proposition 1 showed that households provide the same labor supply over the long run. This proposition shows that it is constant. The homotheticity outcome is intuitive. As high-skilled and low-skilled provide the same labor supply, the labor income ratio equals the wage ratio. The initial assets ratio being equal to that ratio, the total income and consumption ratios follow the same pattern.

An immediate consequence of this proposition is a unique long-run equilibrium with balanced growth. Indeed, consumption and asset shares of aggregate values for each household class remain constant. Aggregate and average dynamics are thus identical to the representative agent case. Consumption and asset shares can be derived from the average values. King, Plosser, and Rebelo (1988) showed that an economy with representative agents converges to a unique equilibrium with balanced growth and constant labor. The heterogeneous agents case with constant skill premium will thus also converge to a unique equilibrium with balanced growth. Proposition 2 is thus a generalization of the King, Plosser, and Rebelo (1988) result to heterogeneous agents separated by an exogenous constant skill premium: the substitution and income effects cancel out over the long run, labor supply is constant even if real wages increase with technology and there is a unique equilibrium\(^{14}\).

I now show the effect of an unanticipated tax reform.

Proposition 3. (Labor supply curve, labor income tax introduction): consider an economy with one representative agent or several household classes separated by a constant skill premium (\( \alpha = 0 \)) in steady-state equilibrium. Let government introduce without

\(^{14}\)This result has already been provided by García-Peñalosa and Turnovsky (2012).
notice labor income taxes at time \( t_0 \), financing government consumption with no utility value for households. Then under a partial equilibrium evolution (exogenously given wage and interest rates after \( t_0 \)), labor supply after the reform is reduced: for small \( \Delta t > 0 \),

\[
l(t_0) > l(t_0 + \Delta t).
\]

I provide a sketch of the proof for the representative agents case and leave the details and heterogeneous case for the appendix. Before doing so, note that events considered in proposition 3 have similarities with those of lemma 5, if one loosely considers a taxed household as poorer and a non-taxed household, before the reform, as richer. The reform would then be similar to an increase in inequality, as the net wage after the reform is lower. The lemma would then predict a lower supply of the poor, that is, of the taxed households. One can however not rely on this argumentation, as lemma 5 applies to cross-sections and proposition 3 to individual responses to net wage shocks.

**Proof (sketch):** the idea is to compare allocations in three worlds. Household of type 1 lives in an untaxed Ramsey economy, and is the sole representative agent. Household of type 2 lives in an identical economy except there have always been taxes. Household of type 1/2 lives in an identical economy except there are no taxes up to point \( t_0 \) and unanticipated taxes afterward. We are interested in the labor supply decisions of the last type of household, before and after \( t_0 \). Since the reform is unanticipated, clearly

\[
l_{1/2}(t) = l_1(t) \text{ and } c_{1/2}(t) = c_1(t) \text{ for all } t \leq t_0.
\]

I will compare the allocations of households of type 1 and type 2 and derive a comparison for households of type 1/2 and type 2. Usage of optimality conditions from lemma 1 will lead to the result, \( l_{1/2}(t_0 + \Delta t) < l_{1/2}(t_0) \).

Households of types 1 and 2 live in standard neoclassical economies so, over the long run, their labor supply is constant (since \( \theta = 1 \)). Because economies are identical except for the tax rate, households of both types take identical decisions, up to a scaling factor. Consumption decisions are homothetic,

\[
c_2(t) = l_2/l_1(1 - \tau) c_1(t).
\]

Since labor income differ by the same factor and both households start with no assets, assets position over time are also homothetic,

\[
a_2(t) = l_2/l_1(1 - \tau) a_1(t).
\]

With a contradiction hypothesis argument and a Taylor approximation, one can show that the scaling factor satisfies \( l_2/l_1(1 - \tau) < 1 \). One can also show that \( l_2 \geq l_1 \), applying a partial equilibrium version of lemma 2. Consumptions and assets of household type 2 are thus strictly smaller at every point in time, \( c_2 < c_1 \) and \( a_2 < a_1 \). When the reform is enacted, households of type 1/2 have the same assets as households of type 1, larger than what they would have built if they always had lived under a taxed regime. They have thus excess assets \( a_2(t_0) \), which they can afford to consume over time, so \( c_{1/2}(t) > c_2(t) = l_2/l_1(1 - \tau) c_1(t) \) for all \( t > t_0 \).

Relative optimality conditions from lemma 1, in a partial equilibrium setting, also ap-
plies when comparing decisions of households of type 1 and type 1/2, so we have

\[
\frac{c_{1/2}(t_0 + \Delta t)}{c_1(t_0 + \Delta t)} = \left( \frac{l_1}{l_{1/2}(t_0 + \Delta t)} \right)^{e} \left( 1 - \tau \right) w.
\]

With \( l_1 \leq l_2 \) and \( l_{1/2}(t_0) = l_1(t_0) = l_1 \),

\[
\left( \frac{l_{1/2}(t_0)}{l_{1/2}(t_0 + \Delta t)} \right)^{e} = \frac{c_{1/2}(t_0 + \Delta t)}{c_1(t_0 + \Delta t)} \frac{1}{1 - \tau} > \frac{l_2}{l_1} \frac{(1 - \tau)c_1(t_0 + \Delta t)}{c_1(t_0 + \Delta t)} \frac{1}{1 - \tau} = \frac{l_2}{l_1} \geq 1,
\]

showing that \( l_{1/2}(t_0) > l_{1/2}(t_0 + \Delta t) \).

QED.

Note first that the partial equilibrium may not be a bad approximation with logarithmic subutility (\( \theta = 1 \)). Indeed, steady-state labor supply is constant (proposition 2), so optimal consumption decisions are defined as in a standard Ramsey model. When there are taxes, the Hamiltonian is \( H(c, l, a) = u_0(c, l) + \mu (ra + (1 - \tau)wl - c) \) and none of the derivatives \( \partial H / \partial c \) nor \( \partial H / \partial a \) involve \( w \) nor \( l \), so the Euler equation is \( \dot{c} / c = (r - \rho) / \theta \), just as the case without taxes. The steady-state interest rates are thus identical, as well as the capital-effective labor ratio and thus wage rates.

Note also that the same result holds if labor income taxes are already in place but increased or if there is an unanticipated negative productivity shock which translates into a drop in wages\(^\text{15}\), with or without taxes. Conversely, if labor income taxes are in place and decrease unexpectedly, households will provide more labor supply.

5 Numerical illustration with representative agents

In this section I make a numerical investigation of labor supply decisions in a standard neoclassical labor market model as in King, Plosser, and Rebelo (1988). The goal is to illustrate theoretical results and show that they are consistent with textbook theory, which considers representative agents. Thus restricting to a model with representative agents, I first illustrate static properties of the labor supply and then dynamic properties. I consider transition paths and steady-states in general equilibrium.

Formally, the model is as in section 3 with a single household class and with labor income taxes. I express it in a recursive fashion in discrete time and solve it by value

\(^{15}\)For instance, with a small open economy assumption. In other partial equilibrium settings, wages are no longer necessarily equal to the marginal product of labor so there can be a decoupling between capital-effective labor ratios and net wages. Other channels than the capital-effective labor ratio have to relay the shock from productivity to net wages.
function iteration over a grid:

\[
V(a_t, A_t) = \max_{c_t, l_t} u_0(c_t, l_t) + \beta V(a_{t+1}, A_{t+1})
\]

s.t. \[a_{t+1} = (1 - \delta) a_t + y_t - c_t \]
\[y_t = (1 - \tau) w_t l_t + f' \left( \frac{a_t}{A_t l_t} \right) a_t \]
\[A_{t+1} = (1 + g) A_t,\]

where \(\beta\) is the time discount factor. I choose the utility function \(u_0(c_t, l_t) = u(c_t) - v(l_t)\) such that it yields a balanced growth path in the steady state, taking \(u(c) = \ln(c)\) and \(v(l) = \lambda l^{1+\varepsilon} / (1 + \varepsilon)\). The production function \(f\) is of Cobb-Douglas form and technology \(A\) grows at exogenously given rate \(g\). Parameter values are chosen to be realistic and have visible effects\(^{16}\). One outcome of the simulation are policy functions \(c^*(a, A)\) and \(l^*(a, A)\) which give the consumption and labor supply choices of the household assuming his assets are \(a\) and the technology level is \(A\).

Results of the experiments are presented in figures 4 and 5, related to the static and dynamic properties of labor supply, respectively. Figure 4 provides a property of the labor supply of representative agents not considered in the paper so far. It illustrates a link with the textbook theory. Figure 5 illustrates the theoretical analysis contained in proposition 3.

![Figure 4: Simulated labor supply curve of a representative agent (smoothed)](image)

The left part of figure 4 plots the labor supply choice of a representative agent depending on its asset level \(a\), for a given level of technology \(A_0\). In other words, it plots the function \(l^*(a, A_0)\). The right part derives the wage level \(w = A_0 (1 - \nu) f(a/A_0 l^*)\) for each asset level \(a\) and corresponding labor choice \(l^*(a, A_0)\) and plots the resulting labor supply curve\(^{17}\).

\(^{16}\)I consider 10 periods, each of length 10 years. The time is discounted with factor \(\beta = 0.96\) (equivalent to a yearly discount of 0.995), the labor disutility elasticity parameter is \(\varepsilon = 0.05\) (giving a Frisch elasticity which is about 5 times larger than the one used by Prescott, 2004), the weight of the disutility of labor component is \(\lambda = 1\), the labor income tax rate is \(\tau = 0.3\), the capital share in production is \(\nu = 0.33\), productivity grows at exogenous rate \(g = 0.3\) (equivalent to a yearly increase of 2.5%), capital depreciates at rate \(\delta = 0.06\) (yielding a steady-state yearly net interest rate of 3.8%).

\(^{17}\)Both curves are smoothed by moving averages, to remove the steps in the plot which come from
5. NUMERICAL ILLUSTRATION WITH REPRESENTATIVE AGENTS

The graphs show that, for low level of assets labor supply is increasing with assets, while it decreases for high level of assets. Correspondingly, the slope of the curve is positive for low wage levels and negative for high wage levels. This plot is consistent with textbook neoclassical labor supply, whose static behavior is recalled in section 3 and represented in figure 2. The difference is that the model here is dynamic.

![Figure 5: Simulated labor supply paths depending on initial assets](image)

Figure 5 plots the labor supply choices of a representative agent over time, depending on its initial asset level \( a_0 \). In other words, it plots the values of the function \( l_*(a_0, A_0) \), \( l_*(a_1, A_1) \), ..., \( l_*(a_{10}, A_{10}) \) for different initial levels \( a_0 \), where \( a_{t+1} = (1 - \delta) a_t + y_t - c_t \) and \( A_{t+1} = (1 + g) A_t \).

The graph exhibits two phenomenons. First, for the chosen asset values \( a_0 \), initial labor supply is higher for low asset levels. Second, the differences disappear over time as labor supply converge to a unique value. These two phenomenons illustrate theoretical results. The second one is consistent with the result from King, Plosser, and Rebelo (1988) that labor supply in steady-state is a well-defined and constant value. The first phenomenon illustrates the transitory wealth effect behind proposition 3. Agents submitted to unanticipated labor tax increases have too much assets at the time of the reform, compared to the new steady state level. On the graph, this would correspond to one of the labor supply path at the bottom (for instance, \( a_0 = 5 \)). When the reform is introduced, they provide a lower labor supply. This is the wealth effect on labor supply. Over time, these differences disappear, showing that the wealth effect is transitory.

Compared to proposition 3, there is one difference. The numerical illustration takes place in general equilibrium, not only in partial equilibrium. The simulation thus provides additional comfort that the partial equilibrium assumption in the proposition comes at a low price.
6 Historical patterns on labor supply in the cross-section

This paper has a theoretical focus and does not contain causal empirical analysis. Confrontation of the theory and broad historical facts however helps to illustrate the potential of the theory. Below I compare in qualitative terms the theoretical testable implications with some stylized evidence on education, skill premiums and the labor market over the 20th century in the United States. Theory and historical empirical patterns appear consistent. In particular, the theory qualitatively explains the US labor supply cross-section evidence over time presented in figure 1.

Figure 6: Enrollment or graduation rates (in %) in the US, 3 education levels

The history of US education is a succession of three transformations (Goldin, 1999a, 1999b). Elementary school was established in the 19th century. Secondary and high school grew at a rapid pace during the first half of the 20th century, adding a practical dimension suitable for professional life. Higher education and university have been developing at a more moderate pace since WWII. Enrollment and graduation figures best illustrate the changes (US Department of Education, 1993). By 1900, essentially all 5 to 13 years old children have been attending elementary school. Less than 7% of the 17 years old graduated from high school in 1900 but more than 50% in 1940 and the number has stabilized at 70%-75% since 1960. Enrollment in higher education institutions grew from less than 3% of the 19 to 24 years old in 1900 to 10% by the end of the WWII. Ever since, it has continued increasing at higher and regular pace, reaching 50% by 1990. Figure 6 presents the evolution between 1900 and 1990. In short, the high school rise is an intense pre-WWII 20th century phenomenon while higher education is a post-WWII phenomenon.

If we compare low and high wage workers over the 20th century, we would expect that the skill premium declined up to WWII and increased after. The transformation of the
education system indeed gave increasing access to secondary school to the population, including low and mid-wage workers. Education allowed this category of workers to close the skill gap. Higher education, of most benefit to high-skilled workers, on the other hand only increased significantly after WWII. It is only then that the skill premium started to increase, at the same moderate pace as enrollment in tertiary education increased. Figure 7 displays the wage premium of college over high school graduates. To the extent that education and skill premiums are reflected in wage premiums, it confirms the expectation that skill and wage premiums declined up to WWII and increased afterward.

Figure 7: College over high school wage premium in the US

Figure 1 displays average daily hours by US males for three dates, for selected wage deciles. It shows that low-wage earners were working more at the beginning of the 20th century and still in 1970, but by a smaller margin. In 1991, low-wage earners were working fewer hours.

Table 1 summarizes the confrontation of the theoretical predictions with broad evidence on education and the labor market. The theory predicts that low-skilled workers work more than high-skilled workers until WWII, since the skill premium decreases. As the skill premium slowly increased after WWII, the model predicts first a narrowing gap in the relative labor supply and then, after a while, labor supply of the high-skilled overcoming that of the low-skilled. These predictions are qualitatively consistent with empirical evidence, as reported by Costa (2000) and Heathcote, Storesletten, and Violante (2010) and illustrated in figure 1.

Heathcote, Storesletten, and Violante (2010) go one step further. They show quantitative consistency between their theory and relative labor supply evidence between 1970 and 2000 in the US, among other labor market and consumption evidence. They also consider exogenous variations in the skill premium but add productivity shocks, incomplete markets, gender and endogenous education decisions to a neoclassical model. The

18The figure is identical when plotting all wage deciles. The selection is made for better visualization of the trends.
present paper shows that, qualitatively, exogenous variations in the skill premium alone are sufficient to explain relative labor supply evidence.

Theories based on home production have also been provided to explain part, but not all of the evidence. Rios-Rull (1993) explains why high-skilled households work more, as in the second part of the 20th century but not the first part. Vandenbroucke (2009) explains the opposite, as in the first part of the 20th century but not the second part.

### 7 Concluding remarks

Among other results, this discussion paper shows three things. First, proposition 1 proves that, in standard neoclassical theory over the long run with constant skill premiums, the income effects cancels out with the substitution effect. The slope of the labor supply curve, in its cross-section dimension, is null. Second, there is no inconsistency between cross-section theory and positive empirical elasticities with respect to net wages. An observed decline in labor supply after a tax increase does not necessarily imply that high wage households work more than others. Third, prolonged increases in skill premium, or wage inequality, leads to larger provision of labor supply by high-skilled households, and vice-versa. Predictions are qualitatively consistent with broad
empirical patterns for the US. The analysis can be extended along several directions. One extension is a confrontation with cross-country patterns. Although conclusions vary by data sources, it is generally admitted that wage inequality increased in the US and the UK in the last decades of the 20th century, while it increased only slightly or remained constant in continental Europe (Acemoglu, 2003). As documented in the special issue of the Review of Economic Dynamics (2010), the hours-wage correlation was mostly positive in the US and UK in the decades leading to 2000 and mostly negative in all continental Europe countries considered (Germany, Italy, Spain and Sweden). These patterns appear consistent with the theory (testable implications 1). A complete analysis could confirm it.

One could also consider earnings patterns and compare them to wage patterns, taking into account heterogeneity in labor supply decisions.

Another extension would be a quantitative analysis to measure the contribution of skill premium increases alone in the variation of the hours-wage correlation observed in the US data. Is it sufficient or does one need endogenous education decisions, gender premiums and incomplete markets as in Heathcote, Storesletten, and Violante (2010)?

If the answer is negative, one could consider alternative mechanisms which can explain the increasing hours-wage correlation and extend the quantitative analysis. One such mechanism could be subsistence consumption. Subsistence consumption pushes low-wage earners to provide more labor than dictated by their distaste of labor. With a higher wage, the pressure of subsistence consumption is relaxed and the influence of labor disutility increases, so that workers will not increase much labor supply. There is thus a friction on the substitution effect. For high-wage earners, there is no need to worry about subsistence so the income effect is not constrained. This could explain historical patterns in hours-wage correlation, just as skill premium variations. Real wages increases brought by technological progress help diminish the influence of subsistence consumption, so ease the friction on the substitution effect. Hours-wage correlation would be negative when real wages are low and subsistence consumption effects strong. It would increase and eventually turn positive as real wages increase. This would be broadly consistent with US patterns of increasing hours-wage correlation over the 20th century, starting very negative early in the century and increasing (Costa, 2000) to turn positive by its end (Heathcote, Storesletten, and Violante, 2010).
Appendices
A. PROOFS

Proof (skill and wage premium differentials): the normalization and notation such that the wage rate of high-skilled equals \( w^R = (1 + \alpha^R)w \) and the wage rate of low-skilled equals \( w^P = (1 - \alpha^P)w \), for \( \alpha^R, \alpha^P > 0 \), is the following.

Let \( w = \partial F(K, AL)/\partial L \) be the marginal product of labor, or aggregate wage rate. For a household of class \( R \), rewards on its labor equals the (per capita) marginal product \( w^R = \frac{1}{N^R} \partial Y}{\partial L} = \frac{1}{N^R} \partial F(K, AL)}{\partial L} = \frac{1}{N^R} wN^RA^R = A^Rw. \) Similarly for low-skilled households, \( w^P = A^Pw. \) The measurement (units) of aggregate productivity is free. For small \( \epsilon^R, \epsilon^P > 0 \), I can scale the initial aggregate productivity \( A^R = 1 + \epsilon^R \) and \( A^P = 1 - \epsilon^P. \) Note that the scaling up or down of \( A_0 \) can be made such that either \( \epsilon^R = 0 \) or \( \epsilon^P = 0 \) (but not both), but it is not necessary. Define \( \alpha^R = \epsilon^R. \) I have then \( w^R = A^Rw = (1 + \epsilon^R)w = (1 + \alpha^R)w. \) Similarly, \( w^P = (1 - \alpha^P)w. \) Introducing the notation \( \alpha, \) one has

\[
\begin{align*}
w^R(t) &= (1 + \alpha^R)w(t) \\
w^P(t) &= (1 - \alpha^P)w(t)
\end{align*}
\]

QED.

Proof (lemma 2): the results also hold when preferences have power parameter no less than unity \( (\theta \geq 1) \) and initial asset distribution satisfy \( a^R_0 \geq (1 + \alpha)a^P_0. \)

Let \( (c^R_*, l^R_*) \) and \( (c^P_*, l^P_*) \) be the optimal allocations chosen by the rich and the poor, respectively. Assume by contradiction that \( l^R_* > l^P_* \). We will build a strict Pareto-improving feasible allocation to prove that we must have \( l^R_* \leq l^P_* \). This allocation does not need to be optimal, simplifying the analysis.

Consider the two following allocations:

\[
\begin{align*}
(c^R_1, l^R_1, c^P_1, l^P_1) &= (c^R_*, l^R_*, c^P_*, l^P_*) \\
(c^R_2, l^R_2, c^P_2, l^P_2) &= ((1 + \alpha) c^P_* - \alpha a^P_*, l^P_*, c^P_*, l^P_*)
\end{align*}
\]

Clearly these allocations leave the poor indifferent. Let us show that they are both improvements for the rich. Because consumption is identical and labor supply \( l^R_1 = l^P_1 < l^R_* \), the first allocation \( (c^R_1, l^R_1, c^P_1, l^P_1) \) decreases labor disutility and is thus an improvement for the rich. By lemma 1, the assumption that \( \theta \geq 1 \) and the contradiction hypothesis, we have

\[
\frac{c^R_{l^R_*}}{c^R_*} = \left( \frac{l^R_*}{l^R_1} \right) \frac{w^R}{w^P} \frac{1}{(1 + \alpha)} < (1 + \alpha) \frac{1}{(1 + \alpha)} \right) < (1 + \alpha)^\frac{1}{(1 + \alpha)}
\]

so, with \( \alpha < 0, c^R_* < (1 + \alpha)c^P_* \leq (1 + \alpha)c^P_* - \alpha a^P_* = c^P_2. \) It follows that the second allocation \( (c^R_2, l^R_2, c^P_2, l^P_2) \) is also a Pareto improvement, as both the consumption utility
increase and the labor disutility component decrease for the rich.

By construction, the allocation \((c^R_i, l^R_i, c^P_i, l^P_i)\) is feasible if \((c^R_i, l^R_i, c^P_i, l^P_i)\) is feasible for an identical factor price schedule \((w^R_i, w^P_i, r_1) = (w^R_2, w^P_2, r_2)\) or for a dominating price schedule \((w^R_1, w^P_1, r_1) \geq (w^R_2, w^P_2, r_2)\), since the consumption of the rich \(c^R_i\) is greater than \(c^P_i\). The latter fact is crucial and comes directly from the contradiction hypothesis that \(l^R_i > l^P_i\). We use the notation \((w^R_i, w^P_i, r_1) = (w^R_2, w^P_2, r_2)\) either in a strict sense, component-by-component sense, or as a short cut for the following: \(w^R_1 \geq w^i_1\) if changes in interest rates are negligible compared to changes in wages. The latter condition is sufficient for the result.

The approach is to show that \((c^R_i, l^R_i, c^P_i, l^P_i)\) is feasible for a selected price schedule \((w^R_2, w^P_2, r_2)\) and verify that \((c^R_i, l^R_i, c^P_i, l^P_i)\) is feasible for the general equilibrium price schedule \((w^R_1, w^P_1, r_1)\) which is consistent with the allocation \((c^R_i, l^R_i, c^P_i, l^P_i)\). We will use this two steps approach because we do not know the price schedule \((w^R_2, w^P_2, r_2)\) which is consistent in general equilibrium with the allocation \((c^R_2, l^R_2, c^P_2, l^P_2)\), but we can demonstrate that it is feasible. On the other hand, we can not directly demonstrate that the allocation \((c^R_i, l^R_i, c^P_i, l^P_i)\) is feasible but we can characterize the price schedule \((w^R_1, w^P_1, r_1)\) in equilibrium. By putting the two together, we will have demonstrated that Pareto improving, this will contradict the hypothesis that \((c^R_1, l^R_1, c^P_1, l^P_1)\) is feasible under general equilibrium. Because it is Pareto improving, this will contradict the hypothesis that \((c^R_1, l^R_1, c^P_1, l^P_1)\) is optimal and thus prove that we must have \(l^R_i \leq l^P_i\).

Consider \((w^R_2, w^P_2, r_2) = (w^R_*, w^P_*, r_*)\) and let us show that \((c^R_2, l^R_2, c^P_2, l^P_2)\) is feasible for that schedule\(^{19}\). For poor households, neither the allocation nor the price schedule changes, so \((c^P_2, l^P_2) = (c^P_*, l^P_*)\) remains a feasible allocation. For rich households, there is a difference. We have that \((c^P_*, l^P_*)\) is the solution of the poor households maximization problem

\[
\max_{c^P(t), l^P(t)} \int_0^\infty \exp(-\rho t)u_0(c^P(t), l^P(t))dt
\]

s.t.

\[
\dot{a}^P (t) = r_*(t) a^P (t) + w^P(t) l^P(t) - c^P(t)
\]

\[
\lim_{t \to \infty} \left[ a^P(t) \exp \left( - \int_0^{t} r_*(s) ds \right) \right] \geq 0.
\]

and similarly for \((c^R_*, l^R_*)\). By definition, an allocation \((c^i, l^i)\) is feasible for a schedule \((w^i, r)\) if and only if

\[
\dot{a}^i (t) = r(t) a^i (t) + w^i(t) l^i(t) - c^i(t)
\]

\[
\lim_{t \to \infty} \left[ a^i(t) \exp \left( - \int_0^{t} r(s) ds \right) \right] \geq 0.
\]

\(^{19}\)Note that in a partial equilibrium setting, with exogenously given wage and interest rates (whether constant or not), it is sufficient to prove this first step only. Indeed, in this case, \((w^R_2, w^P_2, r_2) = (w^R_*, w^P_*, r_*)\) is the equilibrium schedule for both the \((c^R_2, l^R_2, c^P_2, l^P_2)\) and the \((c^R_*, l^R_*, c^P_*, l^P_*)\) allocations. The result holds at any point in time (not only over the long run).
By \((c_2^p, l_2^p) = (c_*^p, l_*^p)\) feasible for the schedule \((w_2^R, w_2^P, r_2) = (w_*^R, w_*^P, r_*)\) we have

\[
 a^p(t) = r_*(t)a^p(t) + w_2^P(t)l_2^p(t) - c_*^p(t)
\]

\[
 \lim_{t \to \infty} \left[ a^p(t) \exp \left( - \int_0^t r(s) ds \right) \right] \geq 0.
\]

Multiplying the budget constraint and the no-Ponzi constraint by \((1 + \alpha)\), we have

\[
 (1 + \alpha)a^p(t) = r_*(t)(1 + \alpha)a^p(t) + (1 + \alpha)w_2^P(t)(l_2^p(t) - (1 + \alpha)c_*^p(t)
\]

\[
 \lim_{t \to \infty} \left[ (1 + \alpha)a^p(t) \exp \left( - \int_0^t r(s) ds \right) \right] \geq 0.
\]

Adding \(\alpha(t)a^p(t)\) on both sides of the budget constraint leads to

\[
 (1 + \alpha)a^p(t) + \alpha(t)a^p(t)
\]

\[
 = r_*(t)(1 + \alpha)a^p(t) + (1 + \alpha)w_2^P(t)(l_2^p(t) - [(1 + \alpha)c_*^p(t) - \alpha(t)a^p(t)]
\]

Define \(a^R(t) = (1 + \alpha)a^p(t)\), consistent with the assumption that initial assets satisfy \(a_0^R \geq (1 + \alpha)a_0^P\). Then \(\partial a^R(t)/\partial t = (1 + \alpha)\partial a^p(t)/\partial t + a^p(t)\partial \alpha(t)/\partial t\). By definition of the skill premium, we also have \(w_2^R(t) = (1 + \alpha)w_2^P(t)\). By definition of the allocation, \((1 + \alpha)c_*^p(t) - \alpha(t)a^p(t) = c_2^p(t)\) and \(l_2^p(t) = l_2^R(t)\). Hence, we have

\[
 a^R(t) = r_*(t)a^R(t) + w_2^R(t)l_2^R(t) - c_2^R(t)
\]

\[
 \lim_{t \to \infty} \left[ a^R(t) \exp \left( - \int_0^t r(s) ds \right) \right] \geq 0.
\]

In other words, the allocation \((c_2^R, l_2^R)\) is feasible for the price schedule \((w_2^R, w_2^P, r_2) = (w_*^R, w_*^P, r_*)\). We have thus shown that the allocation \((c_2^R, l_2^R, c_*^P, l_*^P)\) is feasible for the schedule \((w_2^R, w_2^P, r_2) = (w_*^R, w_*^P, r_*)\).

What remains to be shown is that the allocation \((c_1^R, l_1^R, c_*^P, l_*^P)\) is feasible for the general equilibrium price schedule \((w_1^R, w_1^P, r_1)\). As argued earlier, the allocation is feasible if the price schedule satisfies \((w_1^R, w_1^P, r_1) \geq (w_*^R, w_*^P, r_*)\). Let us characterize the general equilibrium price schedule \((w_1^R, w_1^P, r_1)\) and show it satisfies the inequality.

I start by showing that over the long run the capital-effective labor ratio is higher, \(k_1 = K_1/(AL_1) > k_2 = K_2/(AL_2)\). We have that the aggregate labor supply declines, as, by contradiction hypothesis,

\[
 L_1 = N_1^R l_1^R + N_1^P l_1^P
\]

\[
 = N_1^R l_*^R + N_1^P l_*^P
\]

\[
 < N_2^R l_*^R + N_2^P l_*^P = L_2.
\]

Because rich households work less but consume the same, the aggregate capital stock can also decrease. I need thus to compare explicitly the numerator and denominator of the capital-effective labor ratio.
When relative premiums vary over time (\( \dot{\alpha} \neq 0 \)), there could be borrowing between \( R \) and \( P \) households. In this case, aggregate capital stock \( K < N^R a^R + N^P a^P \). I start with the case of no borrowing between households and return to the other case at the end of the proof.

We have

\[
\begin{align*}
k_1 &= \frac{K_1}{AL_1} \\
    &= \frac{1}{A} \frac{N^R a^R_1 + N^P a^P_1}{N^R I^R_1 + N^P I^P_1} \\
    &= \frac{\frac{1}{A} N^R a^R_1 + \frac{1}{A} N^P a^P_1}{N^R I^R_1 - N^R (I^R_1 - I^R_1) + N^P I^P_1} \\
    &= \frac{\frac{1}{A} N^R a^R_1 + \frac{1}{A} N^P a^P_1}{L^*_s - N^R(I^R_1 - I^R_1)} \\
    &= \frac{\frac{1}{A} N^R (a^R_r - r_1(a^R_r - a^R_1) - w_1^R(I^R_1 - I^R_1))}{L^*_s - N^R(I^R_1 - I^R_1)} \\
    &= \frac{1}{A} K^*_s - N^R w^R_1(I^R_1 - I^R_1) \\
    &= \frac{1}{A} K^*_s - N^R w^R_1(I^R_1 - I^R_1) \\
    &= \frac{1}{A} K^*_s - N^R w^R_1(I^R_1 - I^R_1) \\
    &> \frac{\frac{1}{A} K^*_s}{L^*_s} \\
    &= k_s
\end{align*}
\]

The approximation is based on a first degree Taylor approximation, using the budget constraint, and comes from the following. Consumption is unchanged \( (c^*_i = c^i_1) \) so variations in the accumulation of savings \( (a^*_i - a^i_1) \) are driven by changes in labor and capital income \( (-r_1(a^*_s - a^i_1) - w_1(I^*_s - I^i_1)) \). Over the long run \( (A \to \infty) \), interest rate \( (r_1) \) and asset differentials \( (a^*_s - a^i_1) \) remain bounded (so \( \frac{1}{A} r_1(a^*_s - a^i_1) \to 0 \)) while wages \( (w_1) \) continue to grow, so the approximation of the new asset position \( a^*_i \) is dominated by changes in labor income \( (-w_1(I^*_s - I^i_1)) \), added to the initial equilibrium level \( a^i_1 \).

Note that the production function is Cobb-Douglas so \( w^R_1 = A(1 + \alpha^R)(1 - v) f(k_1) \) (see below). The last inequality comes from the following. Let \( x = N^R(I^R_1 - I^R_1) \). Assuming
that \( k \geq 1 \) and using that the production function is Cobb-Douglas, we have

\[
\frac{\partial}{\partial x} \left( \frac{1}{K} \frac{(1 + \alpha^R)(1 - \nu)f(k)x}{L - x} \right) = \frac{-(1 + \alpha^R)(1 - \nu)f(k)(L - x) + \left( \frac{1}{K} - (1 + \alpha^R)(1 - \nu)f(k)x \right)}{(L - x)^2} \]

\[
\frac{-(1 + \alpha^R)(1 - \nu)f(k)(L - x) + \frac{1}{K}f(k)k}{(L - x)^2} = \frac{k - (1 + \alpha^R)(1 - \nu)k}{(L - x)^2} > 0 \]

The last inequality holds trivially when \( \alpha^R = 0 \) or it is small, since \( k \geq 1 \) and \( \nu < 1 \). As noted in the proof of skill and wage premium differentials at the start of appendix A, one can choose productivity measurement units so that either \( \alpha^R = 0 \) or \( \alpha^P = 0 \) (but not both). The technical assumption of a small \( \alpha^R \) value was made, so the inequality holds.

Hence, the function \( \frac{\frac{1}{K} - (1 + \alpha^R)(1 - \nu)f(k)x}{L - x} \) is growing strictly, \( \frac{\frac{1}{K} - (1 + \alpha^R)(1 - \nu)f(k)x}{L - x} > \frac{\frac{1}{K} - 0}{L - 0} = \frac{1}{L} \) for all \( x > 0 \) and in particular \( x = \frac{\nu R}{1 - \nu} > 0 \), which holds by contradiction hypothesis.

I conclude by showing that the general equilibrium price schedule satisfies \((w_1^R, w_1^p, r_1) \geq (w_2^R, w_2^p, r_2) = (w_*^R, w_*^p, r_*)\). Over the long run, the capital-effective labor ratio is higher, \( k_1 > k_* \). The wage rate \( w = A(1 - \nu)f(k) \) increases with the capital-effective labor ratio, so \( w_1^i = (1 + \alpha^i)w_1 > (1 + \alpha^i)w_* = w_*^i \). The interest rate on the other hand declines. On the long run however, changes in the interest rate are dominated by changes in the wage rate, as verified below. Strictly speaking, we did not show that \( r_1 \geq r_* \), but this is not needed: as discussed above, it is sufficient to show that changes in interest rates are negligible compared to changes in wage rates.

With a Cobb-Douglas production function, \( f(k) = k^v \), \( f'(k) = vk^{v-1} = vf(k)/k \) and \( f''(k) = -v(1-v)k^{v-2} = -v f'(k)/k \), so we have \( w = A(1 - \nu)f(k) \) and \( r = f'(k) - \delta = vf(k)/k - \delta \). The impacts of a change in factor prices from a change in aggregate labor supply \( L \) are, by the envelope theorem, given by \( \partial w(k)/\partial L = w'(k)\partial k/\partial L = A(1 - \nu)f'(k)\partial k/\partial L \) and \( \partial r(k)/\partial L = r'(k)\partial k/\partial L = f''(k)\partial k/\partial L = f''(k)\partial k/\partial L \). Hence, over the long run \((A \to \infty)\)

\[
\frac{\partial r(k)/\partial L}{\partial w(k)/\partial L} = \frac{f''(k)}{A(1 - \nu)f'(k)} = \frac{-(1 - \nu)f'(k)/k}{A(1 - \nu)f'(k)} = -\frac{1}{Ak} \to 0.
\]

In other words, changes in equilibrium factor prices are dominated by changes in wages over the long run. This is not a surprise: in the steady-state of the standard Ramsey model, wages grow at the rate of technology and interest rates are constant.
I finish the proof with the case of borrowing between households, if any incentives and if there are no liquidity constraints which prevent any borrowing. In this case, \( K \neq N^R a^R + N^P a^P \) so the proof that \( k_1 > k_* \) needs to be adjusted.

When the premium declines over some period of time (\( \bar{\alpha} < 0 \) for some time), poor households \( P \) will gain in wage rates over this period, relative to rich households \( R \). Poor have an incentive to borrow from rich now and repay the loans later, when their wage rate is higher in relative terms. Rich have an incentive to make the loan because shifting savings from productive firm capital to other households keep the capital-efficient labor ratio down and the returns on capital high. However, the incentives for between-households borrowing is lower with the \((c_1^R, l_1^R, c_1^P, l_1^P)\) allocations than the \((c_*^R, l_*^R, c_*^P, l_*^P)\) allocations. Indeed, \( l_1^R < l_*^R \) while \( l_1^P = l_*^P \) and wage rates \( w^R \) and \( w^P \) vary in the same proportion, so rich households \( R \) have a lower wage income in the \( l_1 \) case than the \( l_* \) case, relative to poor households. The labor income gap is smaller in the first case than the second case. On the other hand, we saw that wage variations dominate interest rate variations. Overall thus, the income gap between \( R \) and \( P \) is lower with \((c_1^R, l_1^R, c_1^P, l_1^P)\) allocations than \((c_*^R, l_*^R, c_*^P, l_*^P)\) allocations so there are less incentives for between-households borrowing in the first case. Denoting net borrowing flows by \( B_1 \) and \( B_* \) respectively, \( B_1 < B_* \). Then, following the same steps as above,

\[
\begin{align*}
    k_1 &= \frac{K_1}{AL_1} = \frac{1}{A} \frac{N^R a_1^R + N^P a_1^P - B_1}{N^R l_1^R + N^P l_1^P} = \cdots \\
    &> \frac{\frac{1}{A} K_* - \frac{1}{A} B_1}{L_*} \\
    &> \frac{\frac{1}{A} K_* - \frac{1}{A} B_*}{L_*} = k_*.
\end{align*}
\]

QED.

**Proof (lemma 4):** the result also holds when preferences have power parameter no more than unity (\( \theta \leq 1 \)) and initial asset distribution satisfy \( a_0^R \leq (1 + \alpha)a_0^P \).

The proof is almost identical to the proof of lemma 2. For ease of reading, we repeat steps which are non entirely identical. The end of the proof is identical and will not be repeated.

Let \((c_1^R, l_1^R)\) and \((c_*^P, l_*^P)\) be the optimal allocations chosen by the rich and the poor, respectively. Assume by contradiction that \( l_1^R < l_*^P \). We will build a strict Pareto-improving feasible allocation to prove that we must have \( l_1^R \geq l_*^P \). This allocation does not need to be optimal, simplifying the analysis.
Consider the two following allocations:

\[
(c_1^R, l_1^R, c_1^P, l_1^P) = (c_*^R, l_*^R, c_*^P, l_*^P)
\]

\[
(c_2^R, l_2^R, c_2^P, l_2^P) = \left(\frac{c_*^R}{1 + \alpha} + \frac{\dot{\alpha}a^R}{(1 + \alpha)^2}, l_*^R\right)
\]

Clearly these allocations leave the rich indifferent. Let us show that they are both improvements for the poor. Because consumption is identical and labor supply \(l_*^P = l_*^R\), the first allocation \((c_1^R, l_1^R, c_1^P, l_1^P)\) decreases labor disutility and is thus an improvement for the poor. By lemma 1, the assumption that \(\theta \leq 1\) and the contradiction hypothesis, we have

\[
\frac{c_*^R}{c_*^P} = \left[\left(\frac{l_*^P}{l_*^R}\right)^\varepsilon w_*^R\right]^\frac{1}{\theta} = \left[\left(\frac{l_*^P}{l_*^R}\right)^\varepsilon (1 + \alpha)\right]^\frac{1}{\theta} \geq (1 + \alpha)^\frac{1}{\theta} \geq (1 + \alpha),
\]

so \(c_*^R > (1 + \alpha)c_*^P\). With \(\alpha \geq 0\), we then have

\[
c_*^P = \frac{c_*^R}{1 + \alpha} + \frac{\dot{\alpha}a^R}{(1 + \alpha)^2} \geq \frac{c_*^R}{1 + \alpha} > c_*^P.
\]

It follows that the second allocation \((c_2^R, l_2^R, c_2^P, l_2^P)\) is also a Pareto improvement, as both the consumption utility increase and the labor disutility component decrease for the poor.

By construction, the allocation \((c_1^R, l_1^R, c_1^P, l_1^P)\) is feasible if \((c_2^R, l_2^R, c_2^P, l_2^P)\) is feasible for an identical factor price schedule \((w_1^R, w_1^P, r_1) = (w_2^R, w_2^P, r_2)\) or for a dominating price schedule \((w_1^R, w_1^P, r_1) \geq (w_2^R, w_2^P, r_2)\), since the consumption of the poor \(c_*^P\) is greater than \(c_*^P\). The latter fact is crucial and comes directly from the contradiction hypothesis that \(l_*^R < l_*^R\). We use the notation \((w_1^R, w_1^P, r_1) \geq (w_2^R, w_2^P, r_2)\) either in a strict sense, component-by-component sense, or as a short cut for the following: \(w_1^i \geq w_2^i\) if changes in interest rates are negligible compared to changes in wages. The latter condition is sufficient for the result.

The reason to use this two steps approach is identical to the reason in the proof of lemma 2.

Consider \((w_2^R, w_2^P, r_2) = (w_*^R, w_*^P, r_*)\) and let us show that \((c_2^R, l_2^R, c_2^P, l_2^P)\) is feasible for that schedule. For rich households, neither the allocation nor the price schedule changes, so \((c_2^R, l_2^R) = (c_*^R, l_*^R)\) remains a feasible allocation. For poor households, there is a difference. By \((c_2^R, l_2^R) = (c_*^R, l_*^R)\) feasible for the schedule \((w_*^R, w_*^P, r_*)\) we have

\[
ad^R(t) = r_*(t)d^R(t) + w_*^R(t)l_*^R(t) - c_*^R(t)
\]

\[
\lim_{t \to \infty} \left[ad^R(t) \exp \left( - \int_0^t \rho(s)ds \right) \right] \geq 0.
\]
Dividing the budget constraint and the no-Ponzi constraint by \((1 + \alpha)\), we have
\[
(1 + \alpha)^{-1} a^R(t) = r_\ast(t)(1 + \alpha)^{-1} a^R(t) + (1 + \alpha)^{-1} w^R_\ast(t)^{l^R_\ast} - (1 + \alpha)^{-1} c^R_\ast(t)
\]
\[
\lim_{t \to \infty} \left[ (1 + \alpha)^{-1} a^R(t) \exp \left(-\int_0^t r(s)ds \right) \right] \geq 0.
\]
Subtracting \(\alpha(t)a^R(t)(1 + \alpha)^{-2}\) on both sides of the budget constraint leads to
\[
(1 + \alpha)^{-1} a^R(t) - \alpha(t)a^R(t)(1 + \alpha)^{-2} = r_\ast(t)(1 + \alpha)^{-1} a^R(t) + (1 + \alpha)^{-1} w^R_\ast(t)^{l^R_\ast} - \left[ (1 + \alpha)^{-1} c^R_\ast(t) + \alpha(t)a^R(t)(1 + \alpha)^{-2} \right]
\]
Define \(a^P(t) = (1 + \alpha)^{-1} a^R(t)\), consistent with the assumption that initial assets satisfy \(a^R_0 \leq (1 + \alpha)a^P_0\). Then \(\partial a^P(t)/\partial t = (1 + \alpha)^{-1} \partial a^R(t)/\partial t - (1 + \alpha)^{-2} a^R(t) \partial \alpha(t)/\partial t\). By definition of the skill premium, we also have \(w^P_\ast(t) = (1 + \alpha)^{-1} w^R_\ast(t)\). By definition of the allocation, \((1 + \alpha)^{-1} c^R_\ast(t) + \alpha(t)a^R(t)(1 + \alpha)^{-2} = c^P_\ast(t)\) and \(l^R_\ast(t) = l^P_\ast(t)\). Hence, we have
\[
a^P(t) = r_\ast(t)a^P(t) + w^P_\ast(t)^{l^P_\ast(t)} - c^P_\ast(t)
\]
\[
\lim_{t \to \infty} \left[ a^P(t) \exp \left(-\int_0^t r(s)ds \right) \right] \geq 0.
\]
In other words, the allocation \((c^P_\ast, l^P_\ast)\) is feasible for the price schedule \((w^R_2, w^P_2, r_\ast)\). We have thus shown that the allocation \((c^R_\ast, l^R_\ast, c^P_\ast, l^P_\ast)\) is feasible for the schedule \((w^R_2, w^P_2, r_\ast)\).

What remains to be shown is that the allocation \((c^R_\ast, l^R_\ast, c^P_\ast, l^P_\ast)\) is feasible for the general equilibrium price schedule \((w^R_1, w^P_1, r_1)\). This is shown exactly as in the proof of lemma 2 except the indices \(R\) and \(P\) are swapped and \(+\alpha^R\) is replaced by \(-\alpha^P\). The main difference is that \(k - (1 - \alpha^P)(1 - v)k^V > k - (1 - v)k^V > 0\) for all values of \(\alpha^P > 0\), regardless of the normalization of productivity measurement units. If there is any between-households borrowing, \(R\) have an incentive to borrow from \(P\) as \(\alpha > 0\) and the same argument applies, as it is \(P\) households who work less with \((c^R_\ast, l^R_\ast, c^P_\ast, l^P_\ast)\) allocations.

QED.

**Proof (proposition 2):** by proposition 1, we know that \(l^R = l^P\) over the long run. However, we do not know if values are constant. By lemma 1 and \(\theta = 1\), we have
\[
\frac{c^R}{c^P} = \left[ \frac{l^P}{l^R} \right]^{e \frac{w^R}{w^P}} = \frac{w^R}{w^P} = (1 + \alpha).
\]
The law of motion for low-skilled is
\[
a^P = ra^P + w^P l^P - c^P,
\]
while that of high-skilled is
\[
\dot{a}^R = rd^R + w^R l^R - c^R = rd^R + (1 + \alpha) w^P l^P - (1 + \alpha) c^P.
\]

The value \( a^R = (1 + \alpha) a^P \), noting the assumptions \( a_0^R = (1 + \alpha) a_0^P \) and \( \dot{\alpha} = 0 \), represents a solution of the second law of motion. Hence
\[
(1 + \alpha) = \frac{a^R}{a^P} = \frac{rd^R + w^R l^R}{rd^P + w^P l^P} = \frac{y^R}{y^P}.
\]

Consumption and asset shares of aggregate values, for each household class, are thus constant. Aggregate and average dynamics are thus identical to the representative agent case. Consumption and asset shares can be derived from the average values. King, Plosser and Rebelo (1988) showed that an economy with representative agents converges to a unique equilibrium with balanced growth and constant labor supply. The heterogeneous agents case with constant skill premium will thus also converge to a unique equilibrium with balanced growth and constant average labor supply. Since the average labor supply is constant and \( l^R = l^P \), it follows that \( l^R = l^P \) is also constant.

QED.

**Proof (proposition 3):** for a household belonging to class \( i \in \{R, P\} \), let us differentiate tax regimes by an index \( j \in \{1, 1/2, 2\} \) where the index \( j = 1 \) denotes optimal allocations for decisions when there are not taxes, getting net wage \( w_1^i = w^i \); the index \( j = 2 \) the optimal allocations for the household making decisions subject to constant tax rate \( \tau \), getting net wage \( w_2^i = (1 - \tau)w^i \); the index \( j = 1/2 \) the optimal allocations for a household who was not subject to any taxes up to the reform period \( t_0 \) and, without anticipation, is subject to taxes \( \tau \) afterward. The latter household has allocations corresponding to the first household up to time \( t_0 \): \( l_{1/2}^i(t) = l_1^i(t) \) and \( c_{1/2}^i(t) = c_1^i(t) \) for all \( t \leq t_0 \). Let us denote by \( l_{before} = l_1^i(t_0) \) and \( l_{after} = l_{1/2}^i(t_0 + \Delta t) \) for a small time interval \( \Delta t > 0 \) and similarly with \( c_{before} \) and \( c_{after} \). We want to exhibit a transitory wealth effect on labor supply such that labor supply is larger before the reform than after the reform, \( l_{before} > l_{after} \).

First I show that \( l_1^i \leq l_2^i \) in their respective steady states. Lemma 1 also applies, with \( R = 1 \) and \( P = 2 \). The partial equilibrium proof of lemma 2 then also applies and we have \( l_1^i \leq l_2^i \).

Next I compare consumption, incomes and assets of households of type 1 and households of type 2. With \( \theta = 1 \), we know by proposition 2 that labor supply is constant in the steady-state. Denote for tax regime index \( j = 1, 2 \)
\[
\bar{L}^i_j = \int_0^{\infty} \exp(-\rho t) \lambda \frac{(l_j^i)^{1+\theta}}{1+\theta} dt,
\]
a constant number. Then, with preferences \( u_0(c, l) = \ln(c) - \lambda \frac{l^{1+\varepsilon}}{1+\varepsilon}, \) household of type \((i, j)\) solves the following maximization problem, in the steady-state,

\[
\max_{c_j(t), l_j} \int_0^\infty \exp(-\rho t) \ln(c_j(t)) dt - \bar{L}_j
\]

s.t. \( \dot{a}_j(t) = r(t)a_j(t) + w_j(t)l_j - c_j(t) \)

\[
\lim_{t \to \infty} [\dot{a}_j(t) \exp \left( - \int_0^t r(s) ds \right)] \geq 0.
\]

The household skill class \( i \) is fixed and I will compare decisions in different tax regimes \( j \). For ease of notation, I drop the superscript \( i \). Because \( \bar{L}_j \) is constant, the optimal consumption choice \( c_j(t) \) is identical to the choice in a Ramsey model. The Hamiltonian function of the problem is \( \mathcal{H}(c_j(t), l_j, a_j(t), \mu_j(t)) = u_0(c_j(t), l_j) + \mu_j(t) (r(t)a_j(t) + w_j(t)l_j - c_j(t)) \). The optimal solution (see for instance section 8.2, Acemoglu, 2008) is

\[
c_j(t) = c_j(0) \exp \left( (\bar{r}(t) - \rho) t \right)
\]

where \( \bar{r}(t) = \frac{1}{\tau} \int_0^t r(s) ds \) and, for a constant \( M_0 \) identical for both types of households and using the assumption that households have initially no assets,

\[
c_j(0) = M_0 \left[ a_j(0) + \int_0^\infty w_j(t)l_j \exp(-\bar{r}(t)) dt \right] = M_0 \int_0^\infty w_j(t)l_j \exp(-\bar{r}(t)) dt.
\]

Since \( w_2 = (1 - \tau)w_1 \) and labor supplies are constant in the steady-state, we have \( c_2(0) = l_2/l_1(1 - \tau) c_1(0) \) and \( c_2(t) = l_2/l_1(1 - \tau) c_1(t) \). Consumption decisions are homothetic with a factor \( l_2/l_1(1 - \tau) \). Since labor income differ by the same factor, we have incomes \( y_2(t) = l_2/l_1(1 - \tau) y_1(t) \) and assets \( a_2(t) = l_2/l_1(1 - \tau) a_1(t) \).

I next show that \( l_2/l_1(1 - \tau) < 1 \). With a first order Taylor approximation

\[
u_0(c + \Delta c, l + \Delta l) = \ln(c + \Delta c) - \lambda \frac{(l + \Delta l)^{1+\varepsilon}}{1+\varepsilon} \approx \ln(c) - \lambda \frac{l^{1+\varepsilon}}{1+\varepsilon} + \frac{1}{c}\Delta c - \lambda l^\varepsilon \Delta l
\]

\[= u_0(c, l) + \frac{1}{c}\Delta c - \lambda l^\varepsilon \Delta l.
\]

Let \( c = c_1, c + \Delta c = c_2 \) that is \( \Delta c = c_2 - c_1 \) and similarly \( l = l_1 \) and \( \Delta l = l_2 - l_1 \). Assume by contradiction that \( l_2/l_1(1 - \tau) \geq 1 \). Then by homotheticity \( c_2 \geq c_1 \) and \( l_1 \leq (1 - \tau)l_2 < l_2 \) so we have the key consequence that \( \Delta l > 0 \). Let \( \eta \in (0; 1] \) and consider \( l_{\eta} = l_1 + \eta \Delta l \) as well as the consumption level \( c_{\eta} \) which such a labor supply allows in the case of no taxes. By the key consequence \( \Delta l > 0 \), we have a larger labor supply.
A. PROOFS

$l_\eta > l_1$ and in such a partial equilibrium setting, consumption clearly $c_\eta > c_1$. Then

\[ u_0(c_\eta, l_\eta) = u_0(c_\eta, l_1 + \eta \Delta l) \]
\[ \geq u_0(c_\eta, l_1) - \lambda l_1^\varepsilon \eta \Delta l \]
\[ > u_0(c_1, l_1) - \lambda l_1^\varepsilon \eta \Delta l. \]

Because $\lambda l_1^\varepsilon \eta \Delta l$ is bounded and by a continuity argument, one can always choose a small enough $\eta \in (0; 1]$ such that $u_0(c_\eta, l_\eta) > u_0(c_1, l_1)$. We then have found a feasible allocation $(c_\eta, l_\eta)$ which delivers strictly more utility than the optimal allocation $(c_1, l_1)$, a contradiction. Hence $l_2/l_1(1 - \tau) < 1$.

As a consequence, $c_2 < c_1$ and $a_2 < a_1$. When the reform is enacted, at time $t_0$, households have assets $a_1(t_0)$ which are strictly larger than the optimal level of assets $a_2(t_0)$ which they would have built if they had always lived under the tax regime. Shortly after the reform, households have additional assets which they can afford to consume over time. Hence $c_1/2 > c_2$ for some time after the reform.

Using $c_2 = l_2/l_1(1 - \tau)c_1$, we have $c_{1/2} > l_2/l_1(1 - \tau)c_1$. By the partial equilibrium version of lemma 1 applied to households of type 1 versus households of type $1/2$, and using $l_{before} = l_{1/2}(t_0) = l_1(t_0) = l_1(t_0 + \Delta t),$

\[ \frac{c_{1/2}(t_0 + \Delta t)}{c_1(t_0 + \Delta t)} = \left( \frac{l_1}{l_{1/2}(t_0 + \Delta t)} \right)^\varepsilon \frac{(1 - \tau)w}{w} = \left( \frac{l_{before}}{l_{after}} \right)^\varepsilon (1 - \tau) \]

so, with $l_1 \leq l_2$,

\[ \left( \frac{l_{before}}{l_{after}} \right)^\varepsilon = \frac{c_{1/2}(t_0 + \Delta t)}{c_1(t_0 + \Delta t)} \frac{1}{1 - \tau} > l_2 \frac{(1 - \tau)c_1(t_0 + \Delta t)}{l_1} \frac{1}{1 - \tau} = \frac{l_2}{l_1} \geq 1. \]

In other words, we have shown that

\[ \frac{l_{before}}{l_{after}} > 1. \]

QED.
Bibliography


Chapter 3

Time constraints and consumer behavior

Thomas Davoine

March 2011 (updated September 2012, April 2013, August 2013)

Abstract

I take seriously the hypothesis that the wealthy lack time to consume to explain empirical evidence on old age asset decumulation and rich savings rates. Basic life-cycle theory predicts that households run down their assets toward the end of their life but evidence shows they do it at a very low rate. Under homothetic preferences, this theory also predicts that rich and poor save at the same rate, inconsistent with empirical evidence. Integrating a Becker home production model in Ramsey growth theory, I prove that the fact that time is needed for consumption qualitatively explains evidence on savings rate and asset decumulation, as well as secular decreases in per capita labor supply and other evidence.

JEL classification: E21, D9, J14, J22.

Keywords: time constraints, consumption, home production, neoclassical growth theory, savings rate, old age asset decumulation.

1University of St.Gallen (Switzerland) and Institute for Advanced Studies (IHS, Vienna). Email: thomas.davoine@unisg.ch. This paper is an extended version of Davoine (2012). The extension adds no essential results. New sections 2, 3 and 7 expend on the motivation and relevance of the approach. For comments on parts and earlier versions of the paper, I thank C.Keuschnigg, R.Foellmi, A.Kapteyn, J.Mankart, C.Strub, U.Sunde and seminar participants at the University of St.Gallen, the 2011 Annual meeting of the Swiss Society of Economics and Statistics and the 2012 Annual meeting of the Austrian Economic Association. Remaining errors are mine. I am grateful for financial support from the Swiss National Science Foundation and the Risk and Wealth research programme from the University of St.Gallen.
1 Introduction

Standard theory explains the existence of savings by the need to finance consumption after retirement (Modigliani and Brumberg, 1954), to smooth consumption in the face of income changes (Friedman, 1957), because of liquidity constraints (Deaton, 1991) or as self-insurance against uncertainty (Carroll, 1992, 1997). These models can explain many empirical patterns but cannot explain why rich households save at a higher rate and the low rate of asset decumulation in old age. Given the high concentration of wealth distribution in developed economies - 5% of the population holds more than 50% of the assets in the United States - it is interesting to have a good understanding of the saving behavior of the wealthy and how it affects macroeconomic outcomes. I propose a model that is consistent with heterogeneous savings rates and low asset decumulation. The theory can also explain the decline in per capita labor supply in developed economies over the 20th century but does not account for inter-vivo family transfers.

I take seriously the hypothesis that rich households simply lack time to consume and run down their assets. Hurd (1987) concludes that "if one wants to understand how the capital stock is accumulated, one would probably want to study the very wealthy. However, the standard consumption models may not apply: time constraints prevent the very wealthy from consuming even the interest from their wealth." (p.308). In a standard Ramsey growth theory model, I relax the assumption of inelastic labor supply and I am more specific about the consumption process, using a home production approach introduced by Becker (1965). As time is needed for consumption, households allocate a fixed time endowment to either production or consumption. I show that consumption expenditures grow slower than productivity over the long run, because households need to shift time from market production to home leisure activities to enjoy higher consumption. When households are heterogeneous, the theory makes predictions on saving behavior that are qualitatively consistent with evidence.

Following the footsteps of Benhabib, Rogerson and Wright (1991) and Greenwood and Hercowitz (1991), a number of papers have considered home production in macroeconomic analysis. The majority focuses on business cycles or labor market outcomes. Few consider consumption behavior. The model I use can be considered a special case of Vandenbroucke (2009). This paper numerically shows that technological progress raised real wages and made leisure goods more affordable, explaining the decline in per capita working hours in the US between 1900 and 1950, among other labor market outcomes. I extend the investigation, providing analytical results regarding both the labor market and consumer behavior.

The literature on consumption is large. To explain different empirical regularities, a variety of theories have been developed, such as precautionary savings, liquidity constraints, heterogeneous preferences, subsistence consumption, habits, bequest motives or capitalist spirit. Accounting for several empirical patterns at the same time is sometimes a challenge. One contribution of this paper is to account for different patterns of...
1. INTRODUCTION

consumption and labor supply.

Adding time need for consumption helps account simultaneously for the fact that rich save at a higher rate and low asset decumulation in old age. Empirical evidence and firm theory show that workers are not free to choose their hours and that they coordinate labor supply (e.g. Dickens and Lundberg, 1993; Bolton and Dewatripont, 1994). Assuming that heterogeneous households need to coordinate and provide the same labor supply, rich households end up working and earning much. Because time is insufficient for consumption of all their income, rich households save at a high rate, higher than poor households.

Time need for consumption, in a life-cycle partial equilibrium setting, also provides a simple explanation for low asset decumulation in old age, which is especially true for wealthy households (Burbidge and Robb, 1985; De Nardi, French and Jones, 2009): wealthy households simply lack time for consumption over their entire life and end up with assets which they can not run down during retirement.

In addition, the paper qualitatively shows that the theory can explain two empirical findings that are either unexplained or need to rely on the assumption that preferences are heterogeneous, a feature that several authors find unattractive (e.g. Stigler and Becker, 1977). The model is indeed consistent with the evidence that households consumption respond more to predictable future income decreases than increases (asymmetric excess sensitivity; Shea, 1995) and can explain consumption smoothing evidence that appears contradictory in a habit context (Alessie and Teppa, 2010).

In order to test its robustness, I confront the theory to general empirical patterns of consumer behavior. I also show that theoretical predictions are consistent with evidence beyond consumer behavior, looking at various labor supply and macroeconomic evidence. There are a number of empirical patterns not explained by the theory. A more complete picture would require confrontation with empirical analysis that is yet to be performed. The test also provides some ideas for further research on the theoretical front.

In the next section, I provide an overview of the relevant literature and discuss it. Section 3 contains a simple experiment to assess the importance of understanding heterogeneous saving rates. I present the theory under the assumption of a representative agent in section 4. The following section extends the theory to the case of heterogeneous agents. Section 6 compares the theoretical results with empirical evidence and shows in what sense the model helps. Sections 4, 5 and 6 form the core of the analysis. Section 7 extends the confrontation of the theory to more general evidence and the last section concludes.
2 Literature overview and discussion

This section provides an overview of the literature both on theory and empirics, with a focus on consumer behavior and savings rates, a topic of special relevance in this paper. The overview helps put the contribution of the paper into perspective. Readers familiar with the literature can skip this section.

I start with basic life-cycle consumer theory with and without uncertainty, as well as income tracking and asset decumulation patterns.

**Life-cycle theory, income tracking and asset decumulation evidence**  The basic *life-cycle model* of consumer behavior explains the existence of savings by the presence of retirement (Modigliani and Brumberg, 1954): households save while they work in order to finance consumption during retirement. When preferences are concave and intertemporally separable, households will equate marginal utilities of consumption while working and in retirement so that consumption is identical in the two periods. Extending the reasoning to many periods, households will also save to smooth consumption if income changes from period to period, a theoretical prediction known as the *Permanent Income Hypothesis* (Friedman, 1957). In particular, it explains why aggregate and household-level consumption time-series are less volatile than income time-series. The model is however no longer consistent when comparing time-series over longer horizons. Data show that consumption tracks income over the life-cycle and that both are generally hump-shaped (Carroll and Summers, 1991). It also shows that households do no adjust their consumption behavior as soon as new information on future income is obtained as predicted by the hypothesis, but also when new income is actually perceived, an empirical pattern known as *Excess Sensitivity* (Hall, 1978; Campbell and Mankiw, 1990). The model also predicts that savings should be run down during retirement, at odds with data: many households indeed die leaving positive assets behind them.

An unrealistic assumption of this theory is certainty. If the death date is unknown then households may still hold assets when they die, as death can come before it is expected. Uncertainty in future income and imperfect insurance markets may also lead to *precautionary savings* (Carroll, 1992, 1997). This results in smoother consumption, helping to account for excess sensitivity (Deaton, 1992). Since households need some time to build buffer savings early in their career, the model explains the early part of the lifetime consumption tracking pattern of income. It however fails to account for the latter life-cycle part of the pattern, when both consumption and income decline. Another mechanism which explains the early (but not the latter) part of the lifetime income tracking of consumption are *liquidity constraints* (Deaton, 1991). Capital markets are not perfect and young households may not be able to borrow on future expected higher income. On the other hand, *family characteristics* help account for the entire lifetime tracking pattern of consumption, as the hump in consumption takes place at the same time as family size and revenues from work peak (Attanasio and Browning, 1995).
Providing explanations for the various consumption smoothness patterns has occupied the literature for a long time, following the major challenge of the life-cycle model by the excess sensitivity evidence. Precautionary savings, liquidity constraints and, when necessary, family characteristics considerations are considered to be overall successful in reconciling theory with these patterns as well as with old age asset decumulation evidence. At the time of Browning and Lusardi (1996), the consensus in the literature was that these models provide a reasonably fair account of the consumption behavior of low-income and average households, while they can not explain several patterns concerning other households, especially among the wealthy. As Carroll (2000) argued, the rich save at a higher rate than the poor and neither precautionary savings nor liquidity constraints can explain it.

Since then, attention grew on other empirical evidence which relates to all households and is inconsistent with these models. Other than savings rates and wealth distribution patterns, there are bequests evidence and asymmetric patterns. To illustrate the later, Shea (1995) provides evidence of asymmetric excess sensitivity, the fact that households are more responsive to news about future income declines than news about increases.

I now continue the presentation with bequest motive models and various family transfer patterns.

**Bequest motives and family transfers evidence** Once children have left the household, there should be no financial links between them and their parents, according to the life-cycle model. Even though magnitudes are debated, there is however evidence of transfers from parents to their adult children. For instance, Gale and Scholz (1994) estimate that aggregate intervivo transfers amounted to about half aggregate bequests at death in the US in the mid-1980's. Neither the life-cycle, precautionary savings nor liquidity constraints models can explain this evidence. A bequest motive component is needed in the theory.

There are several ways to introduce such a motive in theory. The literature has mostly considered three: pure altruism, paternalistic preferences (or joy-of-giving) and strategic bequests. All three are consistent with evidence of positive assets at death (bequests) and evidence of intervivo transfers, but each is inconsistent with at least one other piece of family transfer evidence. Under pure altruism (Laitner, 1979), parents maximize their own utility as well as that of their children. The latter often have different economic successes in life so transfers should be larger to worse-off children. This prediction is consistent with evidence of compensatory intervivo transfers (Cox, 1990) but not with that of equal bequests at death (Menchik, 1980). Under paternalistic preferences (Andreoni, 1989), parents derive utility from the gifts and bequests they make, without trying to maximize the utility of their children. This could explain why bequests are found to be equal but not why intervivo transfers are compensatory. In the case of strategic bequests (Bernheim, Shleifer and Summers, 1985), parents reward children for services that they provide, such as visits or attention. This variant of the theory could explain
unequal intervivo transfers but not equal bequests at death\(^2\).

I now turn to the presentation of the main theories and empirical patterns relating to saving rates and accumulated wealth.

**Models and evidence on saving rates and wealth accumulation** Under homothetic preferences in a life-cycle model, households with no initial assets and differing only in labor income should have the same saving rate: the consumption and income paths of the richer households are just magnified versions of the paths from poorer households. Dynan, Skinner and Zeldes (2004) show however that households with higher lifetime income save at a higher rate. Figure 1 shows the monotonic increase of saving rates across lifetime income quintiles in the US.

![Figure 1: Saving Rates (%) by Lifetime Income Quintile, US Households, 1984-89](image)

The fact that rich households save at a higher rate should lead to a higher degree of wealth concentration than income concentration, if only accounting (mechanical) effects are at hand. The former can be seen as a stock version and the latter a flow version of the same phenomenon. Quantitatively, this may or may not be the case, depending on which economic forces are at work at the household and macroeconomic level. It is thus instructive to keep account of income and wealth concentration. Diaz-Gimenez, Quadrini and Rios-Rull (1997) for instance calculated that the Gini index of earnings was 0.63 while that of wealth was 0.78 in the US in 1992, confirming that wealth is

\(^2\)Models with infinitely-lived and representative agents, such as the Ramsey model, can be interpreted as bequest motive models with pure altruism. The same holds for models with infinitely-lived agents differing in skills, which will be considered in this paper. Conversely, life-cycle models with pure altruism and perfect inheritance of skills - but not necessarily imperfect inheritance of skills - can be viewed as infinitely-lived agents models.
more concentrated than income. Observation of wealth data reveal several other interesting patterns, sometimes hard to explain by the life-cycle theory and some (but not all) its extensions. For instance, the same authors measure that 7% of US households have zero wealth. Even if rich save at a higher rate, there is no reason for low income households active on the labor market to be isolated from negative shocks and to hold no precautionary savings³.

It is notoriously difficult to build theories which can fully account for wealth concentration in developed economies (Cagetti and De Nardi, 2008).

If the life-cycle model predicts identical saving rates between rich and poor, precautionary savings is even further away from empirical evidence as the model predicts that rich save less. Indeed, rich households can afford larger consumption expenditures so need to save more in absolute terms than poor households to maintain their buffer stock of savings. In comparison to their income however, savings are lower. The liquidity constraints model is also not entirely consistent with the evidence. Indeed, high-income households suffer less restrictions to access capital markets than low-income households. Since some poor households are not able to borrow as much as they wish, in average low income households save more than desired. If anything, liquidity constraints predicts that rich save less than poor.

Bequest motive models are successful at explaining why rich save more, at least under the pure altruism version. The likelihood that children are richer than their parents is larger when parents are poor than when they are rich. When poor, parents have thus little incentive to build a bequest, as intertemporal optimization of consumption over several generations would commend transfers from rich children to poor parents. Thus poor save less. Dynan, Skinner and Zeldes (2004) argue however that it is difficult to understand why rich childless households also save at a higher rate than poor childless households.

Other theories exist, such as heterogeneity in preferences. Rich could simply be the more patient households, explaining why they are richer and why they save more at the same time. If this was the case, one would expect that they patiently build up saving early in life to enjoy higher consumption when they become older. Dynan, Skinner and Zeldes (2004) note however that there is no evidence that rich households decumulate assets at a higher rate when old.

Habits in consumption models lead to the counterfactual prediction that rich save less (Diaz, Pijoan-Mas and Rios-Rull, 2003).

Models with non-homothetic preferences, on the other hand, are consistent with the evidence on savings rates. The need for subsistence consumption, for instance, prevents low-income households to save as much as they would want, absent any constraint. High-income households are little bothered by these constraints and, as a result, save at a higher rate.

³On the other hand, social security may remove the need for private savings by non-participants in the labor market living off welfare benefits.
Under capitalist spirits (Carroll, 2000), households derive utility not only from consumption, but also from wealth itself. Wealth can provide status, can be performance information to ambitious entrepreneurs or capture other motives. Since wealth is more than deferred consumption, there is an incentive to keep accumulating it. At low wealth levels, priority is however given to consumption, in order to survive. The impact of wealth on utility and preferences is larger for rich households, who engage into wealth accumulation and thus save more.

The consumer durables (Fernandez-Villaverde and Krueger, 2011) model is also consistent with the evidence that rich save more. Durable goods, such as real estate, can have a double purpose: consumption and value repository. Early in life, there is an incentive to save via durable goods, to benefit from both purposes. Later in life, once filled with durable goods, households turn to financial assets to save. For savings purposes, durable goods are in average more attractive for low income households and financial assets for high income households. Since saving is measured in financial terms and buying a durable good is only counted as consumption, the saving rate of high income households is higher.

Heterogeneity in saving rates has important consequences. Section 3 will provide a simple example. There is a vast literature which documents and analyzes the consequence of saving behavior on inequality and macroeconomic outcomes. Two early contributions illustrate well the consequence of heterogeneous saving rates. In a neoclassical growth model with exogenous saving rates and identical earnings, Stiglitz (1969) shows that linear or concave savings functions lead to equal distributions of income and wealth over the long run. In the same setting, Bourguignon (1981) shows that convex saving functions (in other words, when rich save at a higher rate) lead to Pareto-optimal unequal distributions, where rich and poor are better off than under equal distributions.

**Labor supply in the cross-section**

In the textbook neoclassical theory, households derive utility from consumer goods and leisure. In its static version, the slope of the labor supply curve is theoretically ambiguous (see for instance Cahuc and Zylberberg, 2004): the substitution effect dominates at low wage levels (positive slope) and the income effect dominates at high wage levels (negative slope). In the dynamic version with representative agents and preferences consistent with balanced growth (King, Plosser and Rebelo, 1988), the substitution and income effects cancel out. Extensions to heterogeneous agents with exogenous and constant skill differences show that high-skilled households provide more or less labor than low-skilled households, depending on initial conditions. Maliar and Maliar (2001) numerically find that high-skilled households work less. Garcia-Peñalosa and Turnovsky (2012) analytically prove that high-skilled households work less if their initial capital endowment, relative to low-skilled, is larger.

---

4Bénabou (1996) and Aghion, Caroli and Garcia-Peñalosa (1999) provide overviews of this literature.
5Chapter 2 of the thesis provides an extensive review of the related literature. I summarize here the parts which are relevant for this chapter.
than their skill premium, and vice-versa.

Adding technological progress and variations in the skill premium, chapter 2 of the thesis proves that high-skilled households provide less labor if the skill premium decreases long enough over time, and vice-versa.

Cross-section evidence on male labor supply shows a shift over the 20th century in the US. Hours-wage correlations were negative at the start of the century but narrowed down over time and turned positive by the end of the century (Costa, 2000; Heathcote, Storesletten and Violante, 2010). Similar evidence are reported for several developed countries over a shorter time horizon, the correlation remaining negative in continental Europe over the entire horizon (Krueger, Perri, Pistaferri, and Violante, 2010).

Two empirical patterns, which do not focus on the cross-section dimension of labor supply, are worth mentioning. First, average labor supply per capita has been decreasing in many developed countries. It decreased in the US before WWII (Vandenbroucke, 2009) and remained constant after. It decreased in most European countries after WWII (Prescott, 2004). Second, empirical analysis comparing labor supply before and after tax reforms find a small but positive elasticity of labor with respect to net wages for males, and larger for females (Blundell and MaCurdy, 1999).

These empirical patterns may appear inconsistent, suggesting different signs for the slope of the labor supply curve. On the one hand, negative hours-wage correlations suggest that the slope of the curve is negative. On the other hand, a positive elasticity of labor supply with respect to net wages suggests a positive slope. Chapter 2 shows that the inconsistency is only apparent. An observed decline in labor supply after a tax increase does not necessarily imply that high-wage households work more than others: a tax increase may lead high- and low-wage households to work less, but it does not prevent low-wage earners to work more than high-wage earners before and after the tax reform.

Chapter 2 presents other extensions of the neoclassical theory from the literature, including inflation, home production, institutions, incomplete insurance markets and imperfect labor markets.

**Home production in macroeconomics** Benhabib, Rogerson and Wright (1991) as well as Greenwood and Hercowitz (1991) add home production into Real Business Cycle models. They show that home production improves the match of theory with empirical fluctuations. Rios-Rull (1993) also considers endogenous education decisions and explains why high-skilled work less at home and more on the market. The analysis by Tanzi and Zee (1993) shows that time constraints on consumption may impact the link between savings and interest rates, of particular importance for econometric evaluations, but contains no macroeconomic result.

Although home production has been considered in different macroeconomic analysis settings, few papers consider long-run consequences and none consider consumer be-
CHAPTER 3. TIME CONSTRAINTS AND CONSUMER BEHAVIOR


3 Simple counterfactual experiment

With a simple experiment I investigate the importance of heterogeneous saving rates in a macroeconomic context. In a partial equilibrium neoclassical growth setting with exogenous constant savings rates and heterogeneous agents, I will compare aggregate output when agents have the same saving rate to output when agents have different saving rates. For increased relevance, I will use actual savings rates observed in the data.

Figure 2 provides another view of the heterogeneity in saving rates, complementary to the one presented in the previous section.

It shows that the top 10% labor income earners in the United States in 1989, when combining their receipts, take home 30% of the total aggregate earnings. The top 10% wealth holders on the other hand own more than 60% of the total wealth. The correlation between income and wealth is not perfect: some wealthy households earn little labor income. There is however a fair link between the two so in this simple example I will assume that a high income earner is a large wealth holder. Under these circumstances, the figure shows that the ratio of earnings to wealth decreases markedly as we move up the top of the distribution, an indication of increasing saving rates.

I now perform the following experiment. I assume that every household behaves the same way and has the same saving rate as households in the middle quintile of the distribution. I then extrapolate a counterfactual aggregate wealth level, take it as capital stock and derive the corresponding output level in a simplified neoclassical growth setting, for comparison with the actual output level. I assume a partial equilibrium setting with constant interest and saving rates as well as unchanged labor supply and earnings. The production function is a Cobb-Douglas function with labor augmenting technology and capital share $\nu = 0.36$, as in De Nardi (2004). Technological progress is unaffected and
productivity grows at an annual rate of 2%, equal to its average between 1950 and 2008 in the US.

Let $W_i$ be the actual wealth of an income group $i$ and $\hat{W}_i$ be its counterfactual level, and similarly for the capital stock $K = \sum_i W_i$ and $\hat{K} = \sum_i \hat{W}_i$. Under the assumptions that I made, group earnings $E_i$, aggregate labor supply $L$ and productivity $A$ remain unchanged. Let also $W^{x,i} = W_i / \sum_u W_u$ be the share of wealth of group $i$ out of aggregate wealth and similarly for earnings, $E^{x,i} = E_i / \sum_u E_u$.

In the simple partial equilibrium setting considered, the ratio of earnings $E_i$ to wealth $W_i$ is a function of the constant saving rate $s_i$, which we can write $E_i / W_i = g(s_i)$. In graphical terms, this ratio divides the size of the earnings bar by that of the wealth bar in the double histogram plot of figure 2. It is an indirect measure of the saving rate $s_i$. Let $j$ be the index for the middle quintile group. The counterfactual experiment consists in applying the saving rate of group $j$ to group $i$ assuming unchanged earnings $E_i$, so

$$\hat{W}_i = \frac{E_i}{g(s_j)} = \frac{E_i}{E_j} W_j.$$

The counterfactual capital stock is thus

$$\hat{K} = \sum_i \hat{W}_i = \sum_i \frac{E_i}{E_j} W_j = \frac{W_j}{E_j} \sum_i E_i = \frac{W_j}{E^{x,j}}.$$

Figure 2: Earnings and Wealth distribution, US, 1989: fraction of total (%) held by top (%).
and the counterfactual output, relative to actual output, is\(^6\)

\[
\frac{\hat{Y}}{Y} = \frac{F(\hat{K}, AL)}{F(K, AL)} = \frac{AL f(\hat{K}/AL)}{AL f(K/AL)} = \left(\frac{\hat{K}}{K}\right)^V = \left(\frac{\hat{K}/AL}{\hat{K}}\right)^V = \left(\frac{W_j/E^{s,j}}{\sum_u W^u}\right)^V = \left(\frac{W^{s,j}}{E^{s,j}}\right)^V.
\]

The data behind figure 2, from De Nardi (2004), contains 72% as the cumulated earnings share of the top 40%, 89% as the cumulated earnings share of the top 60%, 93% the cumulated wealth ownership of the top 40% and 98% the cumulated wealth of the top 60%. The middle quintile group thus has a share of 17% of earnings and 5% of wealth. The counterfactual output level with identical saving rates, compared to the actual output, is thus

\[
\frac{\hat{Y}}{Y} = \left(\frac{5}{17}\right)^{0.36} = 64%.
\]

At an annual growth rate of 2%, the gap between the counterfactual and actual output level corresponds to over 22 years of growth.

This simple experiment illustrates the importance of the heterogeneity of savings rate found in the data. Assuming that all households decide to save at the same rate as the median household from one day to the other and neglecting general equilibrium effects, I find that rich households would save so much less and poor households only a little more such that aggregate capital stocks drops much and that output drops more than 1/3, to a level found 22 years ago. To summarize crudely, neglecting heterogeneity of saving rates cuts more than twenty years of solid two percent annual growth, a third of gross domestic product.

This simple estimate by no means should be attached with any policy implication value. It is a back-of-the-envelope calculation with unrealistic assumptions, including constant earnings and saving rates. In reality, drops in saving rates would be associated with lower capital stock and higher interest rate, triggering a response in saving and thus mitigating the initial cut. General equilibrium effects need to be taken into account for a proper analysis. The example however shows that one can not simply use models which result in identical saving rates across heterogeneous agents, or one may end up with estimates which may be as wrong as one third in value.

---

\(^6\)Households save not only part of their earnings (labor income) but also part of their capital income. Saving rates should thus be expressed in terms of total income rather than earnings. Doing so will however make little difference in the experiment, because the share of earnings and the share of total income from the middle quintile group \(j\) is similar. Earnings and total income indeed differ at the tails of the wealth distribution. Hence \(W_j/E^{s,j}\) is close to \(W_j/Y^{s,j}\) for the middle quintile group with total income \(Y^j\).
4 Theory with a representative agent

4.1 Model

My goal is to remain as close as possible to neoclassical macroeconomic theory to isolate the effect of time constraints on consumption and savings. Essentially, the model is a Ramsey neoclassical growth model where the assumption of inelastic labor supply is dropped and where I am more specific about the consumption process, using the Becker (1965) home production model.

Households: population is constant in size and normalized so that there is a mass 1 of infinitely lived identical households. Households benefit from a time endowment, which is constant, equal for every period of their life and normalized to 1. Households have to make two decisions at every period: since consumption takes time, they have to decide how to allocate time between labor market production and consumption; they also have to decide how much to consume and how much to save from the market production output, a single composite good. Attached to these decisions are the time and goods budget constraints, respectively:

\[ l + (1 - l) \leq 1 \]
\[ \dot{k} = y - c - \delta k, \]

where \( l \) is time dedicated to market production (labor), \( 1 - l \) is time dedicated to consumption activities, \( k \) is the capital stock, \( y \) is output, \( c \) are purchases of the market produced consumer good and \( \delta \) is the capital depreciation rate. Maximizing life-time utility, the time budget constraint will hold with equality and thus become trivial. The goods budget constraint is a standard capital law of motion.

Consumption, home production and preferences: following Becker (1965), households derive utility not from consumer goods themselves but from the combination of consumer goods and the time needed to acquire, prepare and use them. For instance, a bike alone delivers no utility; combining a bike with time, one enjoys making bike tours. Let \( u(c^h) \) be a neoclassical utility function and \( c^h \) be the quantity of the consumer good after acquisition, preparation and usage by a household, which Becker (1965) calls “consumption commodity”. Then the household has instantaneous utility \( U = u(c^h) \). Consistent with the literature, I consider that \( u \) is an isoelastic utility function with parameter \( \theta \). Households discount future utility flows at a rate \( \rho \).

In line with Becker (1965), I assume the existence of a home production function\(^7\)

\[ c^h = h(c, 1 - l). \]

\(^7\)This specification of the home production leaves no room for home productivity increases, for simplicity. It is possible to allow for productivity increases \( A^h \) in \( c^h = h(c, A^h (1 - l)) \) and obtain similar results.
For ease of reading, I will write “home production activities” instead of “home production and consumption activities”, but one should keep in mind that both kinds of activities take place.

In order to obtain analytical results, I put some structure on the home production function. As Greenwood and Hercowitz (1991), I assume that it is a CES function with elasticity parameter $\sigma$, a general functional form that remains tractable:

$$c^h = h(c, 1 - l) = \left( \gamma c^{\frac{\sigma-1}{\sigma}} + (1 - \gamma)(1 - l)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \tag{2}$$

One important assumption of the model is the following:

Assumption H1: the substitution between home production and consumption time $1 - l$ and market produced goods $c$ is difficult or undesirable, so the elasticity $\sigma$ in the home production function is low. Goods and consumption time are complements.

Intuitively, this assumption is plausible: for instance, if a household would like to watch more television but does not have the time to do it, purchasing a more expensive television set will not help.

There are a number of human activities where substitution of time by consumer goods appear to make sense, such as cleaning its own house. There are also a number of other activities where substitution does not appear to make sense, such as watching a movie or playing tennis. For many leisure activities, spending more time is the goal. A 2 hours bike tour is often preferred to a 10 minutes tour. The low elasticity $\sigma$ thus not only captures the feasibility of substitution, but also human desire.

Empirical verification of this assumption is difficult, since the output of home production is not observed yet needed to measure the elasticity $\sigma$. However, indirect evidence is supportive. The literature on optimal indirect taxation justifies lower taxes on food because it is complement to work and higher taxes on luxury goods because they are substitutes to work. In other words, it assumes some goods are complement to leisure. Interpreting an empirical study on recreation activities, Becker (1965) concludes that $\sigma$ is smaller than one. Gronau and Hamermesh (2006) show that highly educated households in the US and in Israel spend more for non-leisure home production activities than low educated households, and are able to save some time from these non-leisure activities (for use in pure leisure or wage earning activities). Time savings are small however.

Table 1 shows that highly educated households spend about 70% more and save about 1% of time in the US. Corresponding numbers are 45% and 2% in Israel. These numbers suggests that $\frac{1}{70} \leq \sigma \leq \frac{2}{45}$, indeed a low value\(^8\).

\(^8\)This estimate of $\sigma$ can not be compared to estimates of the elasticity of substitution between home produced and market produced goods, which several authors have found to be around 2 (for instance, Rupert, Rogerson and Wright, 1995), for two reasons. First, substituting time and goods is not the same as substituting goods from two different production processes. Second, my definition of home production
4. THEORY WITH A REPRESENTATIVE AGENT

### Table 1: Time and consumption expenditures, US and Israel

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>887</td>
<td>1243</td>
</tr>
<tr>
<td>High</td>
<td>879</td>
<td>2124</td>
</tr>
<tr>
<td>Δ</td>
<td>-8</td>
<td>881</td>
</tr>
<tr>
<td>(%)</td>
<td>(-1%)</td>
<td>(+70%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Education</th>
<th>Time</th>
<th>Goods</th>
<th>Time</th>
<th>Goods</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 − l</td>
<td>c</td>
<td></td>
<td>1 − l</td>
<td>c</td>
</tr>
</tbody>
</table>

Market production: consumer goods are produced by market production as in the Ramsey neoclassical growth theory model, with output given by a Cobb-Douglas production function \( y = F(k, A) = k^\nu(A)^{1-\nu} \), factors paid their marginal product, capital stock depreciation rate \( \delta \), capital law of motion \( \dot{k} = y - c - \delta k \) and exogenous technological progress \( \dot{A}/A = g \). For technical reasons, we assume that the capital share parameter \( \nu \) is high enough that \( \nu > (r + \delta + (1 - 1/\sigma)g)/(g + \delta) \), a condition which holds for developed countries.

Labor market: there are two reasons to drop the assumption of inelastic labor supply from the Ramsey model. First, data show that labor supply has decreased over centuries. Second, households may choose a changing allocation of their time to market (labor \( l \)) and home production (consumption time \( 1 - l \)). It is not realistic for time to be entirely allocated to labor market (households need to sleep and eat to survive) or to consumption (even with a miraculous production technology, one would need to control production output), so I assume \( 0 < l_{\text{min}} \leq l \leq l_{\text{max}} < 1 \).

4.2 Comparison with other models

For a good understanding of the model set up, it is worth pointing out similarities and differences with the literature. In particular, it is useful to discuss the expression which results from combining preferences and consumption process,

\[
   u_0(c, l) \equiv u(h(c, 1-l)) = u \left( \gamma c^{\frac{\sigma-1}{\sigma}} + (1-\gamma) (1-l)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}},
\]

which I call the composite utility function.

Mathematically, this expression is similar and formally equivalent to theories of endogenous labor supply where one uses an aggregation function such as \( h(c, 1-l) \) to activities includes leisure. Hiring external help for non-leisure activities (e.g. home cleaning) may be easy, but it is neither easy nor desirable for leisure activities (e.g. viewing television).
CHAPTER 3. TIME CONSTRAINTS AND CONSUMER BEHAVIOR

compare utility from consumption goods and utility from leisure. The key difference is that \( h(c, 1 - l) \) here is not a preference aggregation function but a production function, which measures the output of the Becker (1965) consumption commodity which can be produced with consumption goods and consumption time. Similarly, time spent out of market production, \( 1 - l \), does not represent leisure but an input into a production process, that of consumption. In itself, non-market time \( 1 - l \) does not provide utility. The set up is the strict implementation of the Becker (1965) idea that consumption goods \( c \) alone do not provide utility, but need to be associated with time \( 1 - l \) to be enjoyed.

In reality, there can be a third usage of time, besides market production and consumption activity. Non-market time \( 1 - l \) can also generate utility when not associated with any consumption good. Meditation is such an example of pure leisure activity. For simplicity, I will not differentiate pure leisure time in my analysis and stick to two usages of time, as Becker (1965). The analysis will already provide non trivial results. One could extend the analysis presented here by following Gronau (1977), which added pure leisure to the Becker (1965) theory. This is left for future research.

The model presented here is not the first one which considers home production in a macroeconomic setting, building on the Becker (1965) theory. In contrast to Benhabib, Rogerson and Wright (1991), Greenwood and Hercowitz (1991) and most of the literature which followed it (see section 2 for references), I follow strictly Becker (1965) in assuming that time and goods are imperfect substitutes in the consumption process. With one exception, this literature explicitly or implicitly assumes easy substitution so that market produced goods can be consumed with little or no time. Those are theories of alternative production processes of the same output, on the market or at home. In the model presented here, it is not possible to consume goods without time.

To the best of my knowledge, the only exception in this literature is Vandenbroucke (2009). In this paper and my set up, time is needed and desired for consumption, as in Becker (1965). Different objectives however lead to other differences in the model set ups. Vandenbroucke (2009) differentiates between general goods and leisure goods, allows for endogenous skill choices and explains detailed labor market outcomes. Leisure goods need time to be consumed, not general goods. Then, the rise in real wages due to technological progress, making (leisure) goods more affordable, explains most of the decrease in the average workweek between 1900 and 1950 in the US: the labor supply dynamics of the model results in an income effect dominating a substitution effect.

The set up of the present paper is a special case of Vandenbroucke (2009) and provides additional results: it removes the difference between general and leisure goods and endogenous skill choices\(^9\), and adds results on consumption behavior outcomes. The calibration performed by Vandenbroucke (2009) results in complementarity between (general) consumption goods and time for consumption. Assumption H1 of the present paper is consistent with this calibration outcome.

\(^9\)The representative agent version of this model is equivalent to the model obtained by setting \( \mu = e = 0 \) in Vandenbroucke (2009).
4.3 Theoretical results

This section will present two main results: as a consequence of time constraints, consumption expenditures will grow slower than productivity past a certain level of wealth and labor supply will decline over time.

Before deriving the results, I provide some intuition and the expected outcomes. Technological progress increases real wages, allowing for more consumption. Because substitution of goods and time is low, households need to shift some of their time endowment from market production to home, to be able to enjoy higher consumption. As a first result, labor supply declines over time. Such a reduction of a production factor will also reduce output so that consumption expenditures will not grow as fast as productivity, a second result which contrasts with the Ramsey steady-state case. In continuation, I proceed to formally establish these two results.

Formally, the household maximization problem is

\[
\max_{c(t), l(t)} \int_0^\infty \exp(-\rho t) u_0(c(t), l(t)) \, dt
\]

subject to

\[
\dot{k}(t) = r(t) k(t) + w(t) l(t) - c(t)
\]

\[
\lim_{t \to \infty} \left[ k(t) \exp\left(-\int_0^t r(s) \, ds\right) \right] \geq 0,
\]

(3)

using the notations \( u_0(c, l) = u(h(c, 1-l)) \) for the composite utility function, \( r \) as the net interest rate and \( w \) as the wage rate\(^{10}\). The first condition is a budget constraint and the second is a standard no-Ponzi condition. One can note that this maximization problem is formally equivalent to that of the Ramsey growth model, except that the utility function \( u \) is replaced by the composite function \( u_0 \) and that households have one more decision variable (namely \( l \))\(^{11}\).

I will be investigating properties of the competitive equilibrium path, the equilibrium values such that households solve their maximization problem, firms maximize profit and markets clear. Under perfect competition, firms maximize profit if and only if factors are paid their net marginal products. Labor markets are perfect so there is no unemployment and the labor market clears by assumption. There is only one type of asset, the capital stock \( k(t) \), so the asset market clears by design as long as the initial value \( k(0) \) equals the given initial capital stock. The economy is closed and the interest rate adjusts to clear the goods market. Formally, the competitive equilibrium path is thus the combination of paths of consumption \( c(t) \), capital stock \( k(t) \), wage rate \( w(t) \) and interest rate \( r(t) \) such that, given an initial capital stock \( k_0 \), they are solutions of the household maximization problem (3), \( w(t) = \partial F(k(t), A(t) l(t))/\partial l(t), r(t) = \partial F(k(t), A(t) l(t))/\partial k(t) - \delta \), budget constraints (1) hold and \( k(0) = k_0 \).

The only deviation from the Ramsey neoclassical growth model is the replacement of

\(^{10}\)For ease of reading, I will drop the time index from the notation.

\(^{11}\)There is a parallel with home production models in short-run macroeconomics analysis: Benhabib, Rogerson and Wright (1991) also reach the conclusion that the household maximization problem is the same as without home production, except preferences are different.
the utility \( u(c(t)) \) by the composite utility \( u_0(c(t), l(t)) \) and the addition of a decision variable, \( l(t) \). There is thus little hope to obtain analytical solutions in a simpler fashion with the model presented here than with the Ramsey model. In particular, there is little hope of obtaining closed-form solutions for the competitive equilibrium path. However, just as with the Ramsey model, it is possible to obtain properties of the competitive equilibrium path. I will not be concerned with replicating exactly properties of the Ramsey model. Instead, I will establish different properties and show that one can derive from them useful testable implications and theoretical insights on the behavior of consumers in the presence of time constraints. The properties will be derived in particular from the first order conditions of the household maximization problem, so I start by specifying these conditions.

The two differences \( u_0 \) and \( l \) in the household maximization problem are sufficient to deliver a different Euler equation than in the Ramsey case. Hamiltonian techniques deliver three first order optimality conditions for the maximization problem:

\[
\frac{\partial u_0}{\partial c} = \mu, \quad \mu_r = \dot{\mu} + \rho \mu \quad \text{and} \quad \frac{\partial u_0}{\partial l} = -\mu w,
\]

where \( \mu \) is the multiplier. Using the CES functional form for \( h \), differentiation of \( u_0 = u \circ h \) and algebra, the two first optimality conditions deliver the following:

**Lemma 1. (Constrained Euler equation):** the consumption path of households is defined by

\[
\frac{\dot{c}}{c} = \frac{1}{\varepsilon_{u_0}(c,l)} (r - \rho) + \frac{\eta_0(c,l)}{\varepsilon_{u_0}(c,l)} \frac{\dot{l}}{l}
\]  

where \( \varepsilon_{u_0} \) is the coefficient of relative risk aversion of the composite utility \( u_0 \) and with:

\[
\varepsilon_{u_0}(c,l) \equiv -\frac{\partial^2 u_0(c,l)}{\partial c \partial l} c = \gamma \left( \frac{h(c,l)}{c} \right)^{\frac{1-\sigma}{\sigma}} \left( \theta - \frac{1}{\sigma} \right) + \frac{1}{\sigma},
\]

\[
\eta_0(c,l) \equiv h(c,l) \frac{1-\sigma}{\sigma} \frac{(1-\gamma)}{(1-l)^{\frac{1}{\sigma}}} \left( \theta - \frac{1}{\sigma} \right) l.
\]

Using the CES functional form of \( h(c,l) \), differentiation and algebra yield the following technical result:

**Lemma 2. (Technical lemma):** under assumption H1, the Euler equation components \( \varepsilon_{u_0} \) and \( \eta_0 \) have the following properties over the long run (that is for large enough \( c \)):

\[
(i) \; \varepsilon_{u_0} > 0, \; \frac{\partial \varepsilon_{u_0}}{\partial c} \geq 0, \; \frac{\partial \varepsilon_{u_0}}{\partial l} \geq 0
\]

\[
(ii) \; \eta_0 \leq 0, \; \frac{\partial \eta_0}{\partial c} \leq 0, \; \frac{\partial \eta_0}{\partial l} < 0
\]

\[
(iii) \; \varepsilon_{u_0}(c,l) = \theta \left( \frac{\partial h}{\partial c} \right) c + \varepsilon_h(c)
\]

where \( \varepsilon_h > 0 \) is the coefficient of relative risk aversion of \( h \) (relative to \( c \)).
This technical result is consistent with intuitive expectations, as the following three remarks illustrate. First, there are no reasons to believe that either $\varepsilon_{u0}$ or $(\partial h/\partial c)c/h$ are constant over time, so $\varepsilon_{u0}$ is unlikely to be constant. Even if labor supply was constant, there is no reason to believe that the dynamics of the economy would be identical to that of the Ramsey model. Second, because $\eta_0 \leq 0$ and $\varepsilon_{u0} > 0$, a decrease in labor supply will have two ceteris paribus effects: increase the capital stock per efficient labor unit, so decrease the interest rate $r$ and thus decrease the first term of the Euler equation (which represents the market production effect); increase the second term of the Euler equation (representing the consumption effect). Which effect dominates is the optimal trade-off question under investigation: how much to reduce labor supply to allow for more consumption but without reducing too much output. Third, because utility is time separable, the coefficient of relative risk aversion $\varepsilon_{u0}$ is the inverse of the elasticity of intertemporal substitution. With constant $l$ and $r$, as consumption expenditure grow, $\varepsilon_{u0}$ will grow (as $\partial \varepsilon_{u0}/\partial c \geq 0$), implying a lower elasticity of intertemporal substitution: households are less and less willing to shift consumption over time to benefit from technological progress, differences between current and future consumption are reduced, meaning that consumption expenditures grow less and less.

Optimality conditions deliver more than the technical lemma. Combining the first and third optimality conditions from the household maximization problem with the Euler equation yield the following conditions on consumption expenditures and labor supply schedules:

$$c = (1 - l) \left( \frac{\gamma}{1 - \gamma} \right)^{\sigma}$$

$$\dot{i} = \frac{1 - l}{l} \frac{1}{\beta(c,l)} \Psi(c,l)$$

where $\beta(c,l) \equiv \frac{1 - l}{l} \frac{\eta_0(c,l)}{\varepsilon_{u0}(c,l)} + 1$ and $\Psi(c,l) \equiv \sigma \frac{\dot{w}}{w} - \frac{1}{\varepsilon_{u0}(c,l)} (r - \rho)$.

I now present the first result, which shows that time constraints on consumption apply to all households, past a certain level of consumption expenditures (thus, wealth). The following technical assumption is needed for the result:

Assumption H2: the economy does not start nor reach a level of capital beyond the golden-rule level so remains dynamically efficient and the discount rate $\rho$ is consistent with the interest rate $r^*$ prevalent when no time is needed for consumption:

$$r - \left( g + \frac{l}{\delta} \right) > 0 \quad \frac{1}{\theta} (r^* - \rho) = g.$$

---

12In the Ramsey model, the Euler equation is $\dot{c}/c = \frac{1}{\theta} (r - \rho)$, with the coefficient of relative risk aversion constant and equal to $\theta$.

13The consumption expenditures expression is just the combination of the first and third Hamiltonian optimality conditions, using the CES form of home production. The labor supply expression is just growth rate differentiation of the consumption expenditures expression, combined with the Euler equation.
Abel, Mankiw, Summers and Zeckhauser (1989) provide empirical evidence that current modern economies are dynamically efficient. The second part of the assumption is consistent with neoclassical growth theory, for instance a standard Ramsey model\textsuperscript{14}.

**Proposition 1. (Consumption growth upper bound):** under assumptions H1 and H2 the rate of growth of consumption expenditures $c$ will be strictly lower than that of technological progress, for large enough consumption values $c$:

\[
\frac{\dot{c}}{c} < \frac{\dot{A}}{A} = g.
\]

The proof of the result is contained in appendix A (full details in Davoine, 2011). To establish the second result, one starts with a growth rate differentiation of the first equation in (5), which establishes a relationship between $\dot{c}/c$ and $\dot{l}/l$. Proposition 1 allows to conclude that labor supply decreases over time (see Davoine, 2011, for details):

**Proposition 2. (Labor supply decline):** under assumptions H1 and H2 households will provide less and less labor on the production market over time, for large enough consumption values $c$:

\[
\frac{\dot{l}}{l} < 0.
\]

The formal analysis does not provide a closed form solution for either consumption expenditures or their growth rate, unlike in the Ramsey model where the growth rate is equal to that of technological progress ($\dot{c}/c = g$) in the steady state. However, the result establishes that consumption growth is lower when time is needed for consumption than when it is not needed (which corresponds to the Ramsey model).

### 4.4 Numerical illustration

Table 2 and figure 3 provide the result of a simulation of household behavior under the hypothesis of time need for consumption. The model is the one described in section 4.1 expressed in a recursive fashion and solved by value function iteration over a grid:

\[
V(k_t, A_t) = \max_{c_t, l_t} u(h(c_t, 1 - l_t)) + \beta V(k_{t+1}, A_{t+1})
\]

\[
s.t. \quad k_{t+1} = (1 - \delta) k_t + y_t - c_t
\]

\[
y_t = w_t l_t + f'\left(\frac{k_t}{A_t l_t}\right) k_t = f\left(\frac{k_t}{A_t l_t}\right) A_t l_t
\]

\[
A_{t+1} = (1 + g) A_t,
\]

\textsuperscript{14}The model without time need for consumption is equivalent to a Ramsey model (households allocate all available time to market production, making labor inelastic).
5. Theory with heterogeneous agents

5.1 Model extension

I present an extension of the model to two skill classes, high and low. There are no training possibilities so agents stay forever in the class given at birth. High skills earn a higher wage and will accumulate more assets, so I refer to this class equivalently as the high-skilled or as the rich. The key question is how heterogeneity changes agents decisions. It is always possible to consume alone but not always to work alone, since some of the production is done in firms. Coordination on the labor market will thus constrain agents decisions in time allocation, while consumption decisions will be unconstrained by the presence of another class.

Population: there are 2 classes of people, separated by a permanent and constant skill premium. For simplicity, I assume perfect substitution in production between high-skilled effective labor and low-skilled effective labor. As a result, high-skilled households earn higher wages $w^R(t)$, a constant deviation from the marginal product of labor $w(t)$, and low-skilled households earn lower wages $w^P(t)$, also by a constant deviation from the marginal product of labor$^{17}$:

---

$^{15}$Time is discounted with factor $\beta = 0.96$, the isoelastic utility coefficient is $\theta = 2$, the elasticity of substitution in the home production function is $\sigma = 0.5$, the weight of consumption in the home production function is $\gamma = 0.8$, the capital share in production is $\alpha = 0.33$, capital depreciates at yearly rate $\delta = 0.06$ and productivity grows at exogenous yearly rate $g = 0.2$.

$^{16}$The average yearly variation rate of variable $X$ in table 2 is defined as the value $i$ such that $X_{75} = (1 + i)^{75}X_0$, where $X_0$ is the initial value of the variable and $X_{75}$ is the value of the variable for the final period, 75 years later.

$^{17}$See chapter 2 of the thesis for details.
Figure 3: Numerical simulation: time paths
\[
\begin{align*}
\frac{w^R(t)}{w^P(t)} &= \left(1 + \alpha^R\right) \frac{w(t)}{w(t)} \\
\frac{w^P(t)}{w^P(t)} &= \left(1 - \alpha^P\right) \frac{w(t)}{w(t)} \\
\iff w^R(t) &= (1 + \alpha) w^P(t) \\
\alpha &\equiv \frac{1 + \alpha^R}{1 - \alpha^P} > 1
\end{align*}
\]

The focus is not to explain the origin of inequality but to consider how differences evolve over time\(^{18}\). The overall population remains constant and its size normalized to 1. Because there are no education possibilities, the fraction of rich and poor households is constant over time, in respective proportion \(0 < N^R, N^P < 1\) (with \(N^R + N^P = 1\)).

**Labor market:** the goal of this paper is not to provide a theory of the labor market but to build a theory of consumption that is consistent with the main characteristics of labor markets: households indeed can consume alone but cannot work alone, except in case of self-employment. Intuitively it is clear that labor supply decisions are not completely free: self-employed workers depend on the expectations of their customers and employed workers need to coordinate with colleagues and management. Lack of labor supply coordination within firms would result in delays and increased costs, so no-coordination firms would be priced out of the market by competition.

More formally, cross-section data show concentration points in the distribution of labor supply, both in the intensive (hours) and extensive (retirement) margins, while the distribution of wages is continuous. Figure 4 illustrates concentration points in hours for US workers in the mid 1980’s. If households use wages (relative price of leisure) to define consumption and labor supply, the continuity of the wage distribution would imply continuity of the labor supply distribution if workers were free in their labor supply choices. This discrepancy between theory and data is sometimes an enigma for labor economists (see for instance, Meghir and Phillips, 2010). There is also direct evidence that households are not free to choose their working hours (Ham, 1982; Kahn and Lang, 1991; Altonji and Paxson, 1992; Dickens and Lundberg, 1993; Steward and Swaffield, 1997). Several reasons have been mentioned to explain why household choices are aligned. There is a strong presumption that social security policy, which often comes with penalties for early retirement\(^{19}\), is responsible for bunching in the retirement decisions (Hurd, 1990; Rust and Phelan, 1997; French, 2005). Coordination costs is another explanation that has been provided to explain labor supply alignment and play an important role in the theory of firms: they can define the characteristics of firms (Becker and Murphy, 1992; Bolton and Dewatripont, 1994) and can explain why workers deviate from their preferences to willingly choose the same hours as co-workers (Weiss, 1996).

To sum up, there are empirical evidence and theories on labor markets and industrial organization which establish that working hours are not freely chosen, in part due to coordination. For the sake of simplicity and clarity of the results, I will assume full

---

\(^{18}\)One can also assume that rich households start with higher initial assets. The same results are obtained. For simplicity of the presentation, I assume that initial assets are identical.

\(^{19}\)For instance there is a 10% tax penalty on 401(k) funds for retirements before age 59 in the US.
coordination of labor supply decisions between workers:

Assumption H3: coordination costs force household classes to agree on the same labor supply, via non-strategic bargaining. If $0 < \gamma < 1$ is the bargaining power of the rich relative to the poor and $l_i^*(t)$ would be the labor supply chosen by household class $i \in \{R,P\}$ absent any constraint, the result of bargaining delivers labor supply

$$l(t) = \gamma l_R^*(t) + (1 - \gamma) l_P^*(t).$$

(6)

This assumption not only simplifies the analysis but is key to the results: rich households are more likely to be time-constrained, so prefer a lower labor supply to be able to enjoy more consumption. Because coordination makes them work more than desired, they generate more income, which they can not consume easily, so their assets keep increasing. Thus rich save more because coordination forces them to work and generate more income than desired and they can not shift enough time for consumption.

5.2 Theoretical results

I will first show in this section that the main results from the representative agent case extend to the case of 2 household classes and then that rich households save proportionally more.

Before the formal analysis, I elaborate on the intuition. High-skilled households earn a higher labor income so can consume more. If there were no coordination constraints on the labor market, they would shift more of their time endowment from market production to home, to enjoy higher consumption. Because of the coordination constraints however, bargaining with low-skilled households forces them to work more than desired: high-skilled prefer to work less hours but are forced to supply more to coordinate in the workplace. They can not shift as much time to consumption as their income would
allow them to do. Complementarity between consumption and time slows down their
consumption growth. Income keeps growing at a high rate and they end up saving more
than desired. The opposite holds for low-skilled households. As a result, high-skilled
households save at higher rate than low-skilled households.

the rich prefer to work less hours than the poor and are then forced to supply more to
coordinate in the workplace

\[ \text{Alone, household } i \in \{R, P\} \text{ would solve the following maximization problem} \]

\[
\max_{\hat{c}_i(t), \hat{l}_i(t)} \int_0^\infty \exp(-\rho t) u_0(\hat{c}_i(t), \hat{l}_i(t))dt \\
\text{s.t. } \frac{d\hat{l}_i(t)}{dt} = r(t) \hat{a}_i(t) + w^l(t) l^R(t) - \hat{c}_i(t) \\
\lim_{t \to \infty} \left[ \hat{a}_i(t) \exp \left( -\int_0^t r(s)ds \right) \right] \geq 0.
\]

The maximization problem is the same as (3) except for the wage schedule. However,
since both households class live in the same economy and because of the assumption
that they coordinate over the labor supply (H3), given the labor schedule \( l(t) \), they each solve:

\[
\max_{c_i(t)} \int_0^\infty \exp(-\rho t) u_0(c_i(t), l(t))dt \\
\text{s.t. } \frac{d\hat{c}_i(t)}{dt} = r(t) \hat{a}_i(t) + w^l(t) l(t) - c_i(t) \\
\lim_{t \to \infty} \left[ \hat{c}_i(t) \exp \left( -\int_0^t r(s)ds \right) \right] \geq 0.
\]

Before engaging into the bargaining over the labor supply \( l(t) \), each household needs to
know what its preferred labor supply is. Therefore, before bargaining, each household
solves the problem \( \hat{P}_i \) and obtains its preferred choice \( l^i(t) \). Because bargaining is non-
strategic, there is no incentive for households to announce a different labor supply \( l^i(t) \) in
the hope of obtaining a better outcome. Households then bargain over the labor supply,
using the allocation rule assumed in H3, to obtain \( l(t) = \gamma l^R(t) + (1 - \gamma) l^P(t) \). They
then take this labor supply as given as well as factor rewards \( r, w \) (noting that \( w^l(t) = (1 \pm \alpha^l_i)w(t) \)) and revise their initial consumption decisions, to obtain their final choice \( c_i(t) \).

Thus, a direct consequence of the assumption that rich and poor households bargain over
labor supply (H3) is the following sequencing outcome:

**Definition 1. (3 stages timing):** under assumption H3, the rich and poor households
follow a three step procedure to make their labor and consumption choices

1. Each household \( i \in \{R, P\} \) separately defines its optimal labor supply \( l^i(t) \), solving \( \hat{P}_i \).

2. Households meet and bargain, leading to the coordinated labor supply \( l(t) = \gamma l^R(t) + (1 - \gamma) l^P(t) \).
3. Households separately revise their consumption decisions, taking the labor schedule \( l(t) \) as given, and obtain optimal consumption plans \( c^i(t) \), solving \( P^i \).

There are two immediate consequences. First, the Euler equation for each household class \( i \in \{R,P\} \) applies at stage 3 and is the same as in the representative agent case. Second, the maximization problem that households solve at stage 1 are identical to problems that households would solve if they lived separately. Results from the representative agent case thus apply. In particular, proposition 2 leads to \( l^i/l^i < 0 \) for \( i \in \{R,P\} \).

By the bargaining assumption H3 and stage 2, \( l = \gamma l^R + (1 - \gamma) l^P \), thus \( l/l < 0 \) over the long run. It is worth recalling these results in the following lemma:

**Lemma 3.** under assumptions H1, H2 and H3 the representative agent Euler equations separately hold for \( i \in \{R,P\} \) and the bargained labor supply schedule \( l \) will be declining over the long run:

\[
\frac{c^i}{c^i} = \frac{1}{\varepsilon_u(c^i,l)} (r - \rho) + \frac{\eta_0(c^i,l)}{\varepsilon_u(c^i,l)} \frac{\dot{l}}{l} < 0.
\]

Other results from the representative agent case extend to the two classes setting. Appendix A shows that

**Proposition 3. (Consumption growth 2 classes):** under assumptions H1, H2 and H3 the rate of growth of consumption expenditures \( c \) for each skill class will be strictly lower than that of technological progress, for large enough consumption values \( c^i \): \( i \in \{R,P\} \)

\[
\frac{c^i}{c^i} \to g + \frac{\dot{l}}{l} < g
\]

Although the growth of consumption expenditures for both classes tend to the same long run value, they are not the same at every point in time. Intuitively, rich have larger resources and thus can afford higher consumption. They suffer more from the time constraints on consumption, which reduce the incentive to postpone consumption today to benefit from high productivity tomorrow. We therefore expect the consumption profile of rich households to be flatter. The proof of the proposition that follows indeed shows that \( c^R/c^R < c^P/c^P \) over the long run. Because households need to coordinate and agree on the same labor supply and because wage rates \( w^i = (1 \pm \alpha^i)w \) grow at the same speed, wage income \( w^i l^i = w^i l \) for both classes grow at the same rate. The same proof also shows that household assets grow at the same rate so total income grow at the same rate over the long run: \( y^R/y^R = y^P/y^P \). Combining and using an additional technical assumption, I formally obtain in appendix A the following result:
Proposition 4. *(Rich save more):* under assumptions H1, H2 and H3, when the speed at which labor changes is bounded so that \( |\dot{l}/l| < (r - \rho) \), rich households will save proportionally more than poor households over the long run:

\[
s^R = \frac{y^R - c^R}{y^R} > \frac{y^P - c^P}{y^P} = s^P
\]

I finish this section with four comments.

First, the proof of the proposition shows more than a higher saving rate from the rich. It shows that the saving rate of the rich grows strictly more (decreases strictly less) than the saving rate of the poor. Indeed, income of the rich and the poor grow at the same rate but consumption of the rich grows strictly less.

Second, the result holds over the long run only. Intertemporal decisions by households could mean that the rich save less, initially. Indeed, the consumption profile of the rich is flatter, while income grows at the same rate for poor and rich. As income grows, the rich anticipate increasing scarcity of time for consumption in the future and could thus shift some consumption from the future to the present. If they shift much consumption, it could be that they initially save at a lower rate than the poor. But such intertemporal shifting of consumption would only increase the saving rate of the rich in the future. The comment above proves that such a lower saving rate by the rich would only be temporary. Further, Dynan, Skinner and Zeldes (2004) provide evidence that rich have been saving at a higher rate in the US in the 1980s, showing that the long-run result is already relevant.

Third, note that the same result holds if labor supply is not only coordinated, but constant: rich save more also if labor supply is constant. The intuition is as follows. Rich earn a higher wage. Because labor supply is constant, their labor income is higher. Technological progress increase future wages and consumption possibilities but constant labor supply prevents from shifting time from production to consumption. The incentive to postpone consumption today to benefit from high productivity tomorrow is thus low, and lower for rich than poor. As in the case of coordinated but variable labor supply, the consumption profile of the rich is flatter over the long run, \( \dot{c}^R/c^R < \dot{c}^P/c^P \). Since labor and the skill premium are constant, earnings of the rich and the poor will grow at the same rate. As in the previous case, total income will grow at the same rate over the long run: \( y^R/y^R = y^P/y^P \). The same conclusion applies, \( s^R > s^P \). The formal proof can be found in Davoine (2011).

The fourth and final comment is an implication of the theory on labor supply preferences: rich households would prefer to supply less labor, if there was no coordination constraint. Indeed, the proof of the proposition 4 shows that \( c^R/c^R < c^P/c^P \) over the long run. Consumption of the poor follows that of the rich with a lag, so at any point in time poor consume less \( c^P_t < c^R_t \). Given the complementarity of goods and time in consumption activities, rich will always want to allocate more time to home (consumption
activities) than the poor, so coordination-free labor time for the rich is lower than the poor, \( l_R^t < l_P^t \).

Section 7 will discuss the consistency of this prediction with empirical evidence. Could this prediction be expected from the onset? I conclude this section with theoretical answers.

Chapter 2 of this thesis investigates the impact of exogenous skill differentials in neoclassical theory, where labor causes disutility, building on the existing literature. One if its result is that the long-run slope of the labor supply curve is null, as substitution effects cancels out with income effects, for constant skill premiums. High-skilled work the same amount of time as low-skilled, in this setting.

If we added a need for time to consume to this theory, one would expect that the slope of the labor supply curve becomes negative. Indeed, the fact that time is needed for consumption, even a tiny bit, introduces a friction in the substitution effect. Because some time is needed for consumption of additional purchases made possible by a higher wage, households will not increase as much labor supply than in a case where no time is needed for consumption. Time constraints on the other hand have no impact on the income effect, as the decrease in the disutility of labor is free. Over the long run thus, the substitution and the income effects will no longer cancel out, when time is needed for consumption. The latter effect will dominate. The slope of the labor supply curve, in a neoclassical model with disutility of labor and time needed for consumption, will be negative over the long run. This chapter shows that one does not even need disutility of labor.

Need for time to consume is one friction in the substitution effect. One can imagine other frictions leading to the same prediction that the slope of the labor supply curve is negative. One such friction are contracts. Beaudry and DiNardo (1995) show that contractual frictions induce a decoupling of wages from productivity, a reason for income effects to dominate substitution effects, so that increases in wages are associated with a decrease in hours. Their theoretical predictions are tested and confirmed. Another possible friction, to be formally analyzed, is subsistence consumption. Subsistence consumption pushes low-wage earners to provide more labor than dictated by their distaste of labor. With a higher wage, the pressure of subsistence consumption is relaxed and the influence of labor disutility increases, so that workers will not increase much labor supply. There is a friction on the substitution effect. For high wage earners, there is no need to worry about subsistence so the income effects is not constrained. Overall, the substitution effect can be dominated by the income effect. The fact that the income effect

\[ \frac{c^R}{c^R} < \frac{c^P}{c^P} \] implies that sooner or later consumption of the poor reaches the level of consumption of the rich, with a lag: \( c^P_{t+\Delta(t)} = c^P_t \) with time-varying \( \Delta(t) > 0 \) so \( c^P_t < c^R_t \). If the coordination-free labor supply choices \( l_P^t \) and \( l_R^t \) resulting from solutions of the problem \( \hat{P}^R \) where such that \( l_P^t \leq l_R^t \), there would be a decision schedule \( (c^R, \bar{l}) \) delivering strictly better outcomes for the rich than the optimal solution \( (c^R, l_R^t) \) to \( \hat{P}^R \). Taking the poor labor supply choices \( \bar{l} = l_P^t \) would deliver a higher composite good level \( c^{h,R} = h(c^R, 1 - \bar{l}) \), as expressed in (2), and so a higher utility level \( u_0(c^{h,R}) \) in the objective function of \( \hat{P}^R \), a contradiction. Hence \( l_P^t > l_R^t \).
6. **Empirical relevance**

I confront theoretical predictions of the consumption time model built in the previous sections with consumer behavior evidence sometimes difficult to rationalize. Section 7 discusses the consistency of the model with other pieces of evidence.

Dynan, Skinner and Zeldes (2004) noted that models with precautionary savings, social security, patience heterogeneity and habits are not consistent with the evidence that rich households save at a higher rate. Capitalist Spirit (Carroll, 2000) and Consumer Durables (Fernandez-Villaverde and Krueger, 2011) on the other hand are consistent with this evidence. Some challenges to these models can motivate one to look for other, complementary, explanations. The consumption time model can be one such complementary explanation.

Specifically, the wealth-in-utility component of preferences in the Capitalist Spirit model is unfortunately not observable and alternative specifications lead to different quantitative predictions: Francis (2009) finds that capitalist spirit preferences increase wealth concentration while Luo and Young (2009) find the opposite with an alternative specification for those preferences. It is difficult to explain with the Consumer Durables model why retired households do not take annuities on their houses (e.g. Chiuri and Japelli, 2010; Poterba, Venti and Wise, 2011) and why the market for reverse mortgages is so small (e.g. Davidoff and Welke, 2007): real estate is an expensive durable consumer good and reverse mortgages would allow old households to make it liquid, if they were interested in consumption and running down their assets.

The theoretical results obtained in sections 4 and 5 apply to a dynastic model. To enable a comparison with the evidence, sometime cast in a life-cycle framework, I show in appendix B that one can derive a life-cycle partial equilibrium version of the consumption time model for which results still apply; that the intuition on consumer behavior is the same; and that the model is suitable for comparison with household-level evidence.

**Savings rate and asset decumulation** In a simple 2 classes setting, the model investigated in this paper showed that time constraints on consumption lead to higher saving rates for rich households (proposition 4).

The model also provides a simple explanation for why assets are not run down by old households, especially wealthy ones. According to the life-cycle model, households should die with no assets if there is no uncertainty on the death date. Yet data shows they die with positive wealth. The consumption time model provides the following explanation: there is so much wealth accumulated at the end of one’s life simply because there was no time to consume it over one’s entire life.
The model is also consistent when looking at asset decumulation evidence for wealthy and non-wealthy households. Burbidge and Robb (1985) find that blue-collar workers decumulate assets after retirement, but not white-collar workers. The first category are less likely to be wealthy than the second category. More directly, data in De Nardi, French and Jones (2009) show that wealthy singles aged 72-81 decumulate assets at a much lower rate than non-wealthy one. The consumption time model with heterogeneous households is consistent with this evidence, since only the wealthy are subject to time constraints.

Beyond evidence on savings rate and asset decumulation, the model is qualitatively consistent with two pieces of evidence from the empirical literature on consumer behavior that are either unexplained or explained with an assumption of heterogeneity in preferences, the latter case being an unattractive feature for some authors (e.g. Stigler and Becker, 1977). Explanations rely on the asymmetric impact that time constraints have on consumption: constraints make it hard for households to increase consumption but not to decrease it.

**Asymmetric excess sensitivity** Confronted with new information about future income, households will adjust their consumption plans. In a basic life-cycle model this adjustment should take place at the time the information is learned (the Permanent Income Hypothesis from Friedman, 1957). Data shows however *Excess Sensitivity*, as households will not fully adjust consumption at the time of the new information but also once income has actually changed (Campbell and Mankiw, 1990). Several theories can explain this pattern\(^{21}\). It is more difficult to account for *Asymmetric Excess Sensitivity* documented by Shea (1995), the fact that households are more responsive to income declines than increases. Loss aversion is a possible explanation for this pattern. Continuing the analysis, Garcia, Lusardi and Ng (1997) established that asymmetric excess sensitivity holds for non-liquidity constrained households, while excess sensitivity is symmetric for liquidity constrained households. In the consumption time model, only the wealthy are unable to freely adjust consumption upward, while they can adjust it freely downward if their income path warrants. Time constraints introduce an asymmetry for the wealthy, who are less likely to be liquidity constrained, not the non-wealthy. In contrast to loss aversion, there is no need to assume different preferences for the two classes of households (only the non-liquidity constrained have aversion for loss).

**Habits, durability or time constraints** Models with habit play an important role as they help explain several puzzling evidence, such as high equity premium (see for instance Campbell and Cochrane, 1999). Finding microdata evidence of habits in consumption has proven a challenge, however. Early microeconometric studies (Dynan, 2000) find no significant evidence of either habits or durability. Recent more sophisticated approach delivered different but intriguing results. Looking for evidence of habits

\(^{21}\)See for instance surveys by Browning and Lusardi (1996) or Attanasio and Weber (2010).
using consumption changes data, Guariglia and Rossi (2002) find the opposite, namely durability; using the same estimation method but savings data, Alessie and Teppa (2010) find evidence of habits and conclude that these two pieces of evidence are at odds with existing theory.

This apparently inconsistent evidence can be explained by the consumption time model, thanks to asymmetries. Indeed, the savings formulation of the empirical Euler equation is $s_{it} = \alpha s_{it-1} + \text{controls}$ and the consumption change formulation is $\Delta c_{it} = \beta \Delta c_{it-1} + \text{controls}$. The empirical findings are $\hat{\alpha} > 0$ and $\hat{\beta} < 0$. In a habit versus durability model, the first evidence exhibits habits and the second durability, leading to an inconsistency between the model and data. In contrast, the consumption time model predicts $\alpha > 0$ and $\beta < 0$, consistently with evidence.

7 General consistency of theory and evidence

This section discusses the consistency of the model with further evidence on consumption and labor supply behavior. I consider here consumption behavior evidence well explained by existing theories, as well as evidence not directly related to consumption.

As will be seen, the model can not account for a number of empirical patterns. The first goal in this section is a robustness test: is the model basis sound enough that it could be modified to account for other empirical patterns? The second goal is an identification of ideas for further research: what are the critical parts of the model to be changed and what other questions (if any) could be addressed with the model?

7.1 Other consumer behavior evidence

I discuss the consistency of the consumption time model with salient empirical patterns of the consumer behavior literature, overviewed in section 2 and not already discussed: various forms of income tracking, wealth distribution, intra-family transfer and childless households behavior. Consistency with evidence on heterogeneous savings rate and old age asset decumulation has already been discussed in section 6. I will also discuss additional selected empirical findings relevant for the model and highlight evidence where it is either silent or inconsistent.

\footnote{Start with savings data: in case of a large positive income shock, a time-constrained household can not easily increase consumption so savings increase more than without time constraints; in case of a large negative shock however, consumption can be freely decreased; so in average, income shocks lead to increases in savings. Hence the prediction that $\alpha > 0$. Continue with consumption changes data: when time-constrained, a household facing a positive shock can not keep increasing consumption as much as he might wish ($0 < \beta < 1$ in this case), but he can have any large decrease of consumption in case of a negative shock ($\beta \leq -1$ in this case); so, in average, current consumption changes are negatively correlated with past consumption changes ($\beta < 0$).}
CHAPTER 3. TIME CONSTRAINTS AND CONSUMER BEHAVIOR

**Income tracking patterns** I start with evidence disregarding wealth heterogeneity. Confronted with new information about future income, households will adjust their consumption plans. Data shows excess sensitivity, as households will not fully adjust consumption at the time of the new information but also once income has actually changed. Data over the entire life cycle shows that the consumption profile tracks the income profile, both hump-shaped. As such, the consumption time model does not explain these empirical patterns. One can however add the family characteristics argument to account for both these evidences (adult-equivalence scale as in Attanasio and Browning, 1995). So far, time endowment was considered constant in the model. In reality, the household’s time endowment is proportional to the number of members in the household. Midway through the life-cycle, there are children in the household, which is thus less limited by time constraints for consumption (children consume but do not work). Hence, the hump-shaped consumption profile.

I consider next the evidence with wealth heterogeneity. As analyzed in details in section 6, the consumption time model can account for asymmetric excess sensitivity and wealth dependent evidence without assuming heterogeneity in preferences. As recalled in Li (2009), empirical research in the 1960’s showed that the consumption of households responds much to small unexpected income (windfall) increases but not so much to large increases. In the consumption time model, it is harder to increase consumption if income increases by much when time is already lacking for consumption. On the other hand, it is easier to consume more if the increase in income is small.

**Wealth distribution patterns** The two basic facts are that wealth is more concentrated than income and that rich households save more. Theoretical results in section 5 explain why rich save more with the fact that labor supply needs coordination and time is short for consumption. Translating the flow result in stock terms leads to larger asset accumulation by the rich and so more concentration in the wealth distribution than in the income distribution, qualitatively consistent with the evidence. Whether this is sufficient to account for wealth distribution patterns is an open question. A quantitative analysis is left for future research.

**Family transfer patterns** The consumption time model is not consistent with family transfer evidence. There is indeed evidence of transfers from parents to adult children, both inter vivos and as bequests. The model presented here is a standard life-cycle theory with the only difference that consumption takes time. There is thus no more motive for inter-family transfer than the standard theory, so the consumption time model fails to account for the evidence.

**Other empirical patterns** Data show a drop in household consumption expenditure after retirement, which appears inconsistent with the consumption time model. Indeed,
7. GENERAL CONSISTENCY OF THEORY AND EVIDENCE

retirement frees up time for consumption and time-constrained households should be in position to increase consumption purchases. However, Aguiar and Hurst (2005) show that the value of total consumption, including home produced goods, does not decline after retirement. Using disaggregated consumption data, Aguiar and Hurst (2008) show that work related expenses account for most of the decline in consumption expenditures after retirement, while expenses remained stable or increased for all other categories of goods. For instance, entertainment expenditures and time increase after retirement. These post-retirement patterns appear consistent with the consumption time model but a more careful empirical analysis is needed.

The model is silent on a number of other empirical evidences, such as the fact that many low income households have no wealth; that households have assets portfolios that appear suboptimal in several respects, such as holding large credit card debts and little equity (Gross and Souleles, 2002); or that savings and wealth are higher for households subject to higher income uncertainty (Carroll, 1997).

7.2 Wealth and labor supply evidence

One theoretical implication of the consumption time model with two agent classes is that wealthy households prefer to work less than other households. Empirical studies confirm this prediction for one labor supply margin. Findings for other margins appear consistent but more empirical research is needed to conclude.

The prediction is consistent with other theories. A standard life-cycle consumer model with heterogeneous agents, uncertainty, borrowing constraints and incomplete insurance markets (Aiyagari, 1994) as well as endogenous labor supply but no coordination also leads to the wealthy working less, as shown by Pijoan-Mas (2006) and Marcet, Obiols-Homs and Weil (2007). Considering infinitely-lived agents who differ by their initial capital endowment in a Ramsey economy with disutility of labor, Turnovsky and Garcia-Penalosa (2008) prove that the wealthy work less because the marginal utility of wealth declines and they prefer to increase consumption of all goods, including leisure. Authors claim that “there is substantial empirical evidence documenting the negative relationship between wealth and labor supply” (p.1401).

The theoretical prediction is confirmed empirically for one labor supply decision margin, retirement. Various empirical studies use unanticipated wealth gains to establish the causality direction and show that increasing wealth leads to earlier retirement. This behavior is found among lottery winners (Imbens, Rubin and Sacerdote, 2001), households blessed with unexpected inheritance (Holtz-Eakin, Joulfaian and Rosen, 1993) or enjoying unexpected stock market gains (Hurd, Reti and Rohwedder, 2009).

For other labor supply margins, there are indications from various empirical studies that the theoretical prediction holds. There is evidence of a negative correlation between wealth and employment probability (Bloemen and Stancanelli, 2001) and that women
provide more labor supply if the household has higher mortgage debt (Fortin, 1995). However, causality in these studies remains unclear and sometimes there is no control for wealth. More systematic empirical analysis is needed for confirmation of the theoretical implication along other labor supply margins than retirement, a topic for further research.

7.3 Skill and labor supply evidence

The theory also implies that rich households would prefer to work less if there were no coordination constraints, because of their higher skill level. This implication is partially but not fully consistent with empirical evidence, similar to the textbook neoclassical theory of labor supply. In this subsection I discuss various theories and the empirical evidence.

Garcia-Peñalosa and Turnovsky (2012) show that the neoclassical theory with constant skill premium differentials predicts that high-skilled households work as much as low-skilled households if the initial wealth endowment differential is equal to that of the skill differential, and less if the endowment differential is higher.

The predictions from the consumption time model and neoclassical theory can appear inconsistent with the small but positive empirical elasticity of labor supply with respect to net wages. As discussed by chapter 2 of this thesis, one can be mislead by comparisons of the cross-section dimension of labor supply with other properties of labor supply. An observed decline in labor supply after a tax increase does not necessarily imply that high wage households work more than others.

There is some but limited cross-section evidence. The prediction that high-skilled work less is consistent with US evidence for the first part of the 20th century but not with the later part (Costa, 2000; Heathcote, Storesletten and Violante, 2010), looking at the intensive labor supply margin of males (hours-wage correlation). For other labor supply margins, there are indications that the theoretical prediction holds: there is evidence of a positive correlation between hourly wage and vacation time (Altonji and Usui, 2007). However, causality remains unclear so more systematic empirical analysis is needed for these margins.

Exogenous changes in the skill premium in a standard neoclassical labor supply model, possibly driven by education changes, can explain hours-wage correlation changes in the US over the 20th century, in broad and qualitative terms. In a richer model, Heathcote, Storesletten and Violante (2010) quantitatively account for changes over the last part of the century (but not average values). Details can be found in chapter 2 of this thesis.

One can expect that the addition of exogenous changes in the skill premium and disutility of labor in the consumption time model would also lead to the high-skilled working more, while maintaining the result that they save at a higher rate. Indeed, high-skilled would still decrease labor supply over time (even though they would do it at a lower
rate than low-skilled, in this case) and coordination would still make high-skilled and low-skilled provide the same labor supply, the two labor market conditions which lead to the result on the savings rate.

7.4 Macroeconomic evidence

In this section I confront the theoretical results with a representative agent from section 4.3 with aggregate macroeconomic evidence. Under the representative agent assumption, one can indeed compare per capita outcomes, observed in the data, with theoretical prediction for the agent behavior.

It is possible to extend the representative agent analysis from consumption to output. In the Ramsey model, not only consumption \( c \) grows at the rate \( g \) of technological progress in the steady-state, but also output \( y \). Proposition 1 proved that \( \dot{c}/c < g \) in the long run and proposition 2 that \( \dot{l}/l < 0 \). The results under time constraints only concern consumption \( c \). Further analysis generalizes this result to output \( y \) and shows that the model is consistent with two of the Kaldor facts of growth over the long run:

**Proposition 5. (Aggregate growth facts):** under assumptions H1, H2, H3 and other technical assumptions, the economy’s competitive equilibrium path is (asymptotically) such that two Kaldor facts of growth hold and output per capita grows slower than technology:

\[
    r = \text{constant} \quad \frac{k}{y} = \text{constant} \quad \frac{\dot{y}}{y} = g + \frac{\dot{l}}{l} < g
\]

Simulations in section 4.4 illustrate the consistency of the aggregate behavior of the model with the two Kaldor facts of growth. The two graphs at the bottom right corner of figure 3 show that the capital per efficiency unit ratio \( k/Al \) (and thus the interest rate) and the capital-output ratio \( k/y \) both converge to a constant value.

Figure 5 plots per capita consumption in efficiency unit \( (C/AN) \) over time and figure 6 per capita labor supply \( (l/N) \) for six countries after the second world war. They identify two challenges of the Ramsey neoclassical growth theory under dynamic efficiency. Specifically, they show that the assumption of inelastic labor supply and the prediction that per capita \( C/N \) grows at least as fast as technology \( (A) \) under dynamic efficiency do not hold, except for the United States\(^{23}\). Figure 5 instead shows that per capita consump-

\(^{23}\)Note that the figure 5 looks different when taking less precise measurements of labor input. The figure is indeed constructed using hours as labor input measure, not the number of workers. Building the same figure using the number of workers as measure for labor, one obtain results which appear more consistent with the Ramsey model assumptions and predictions.
CHAPTER 3. TIME CONSTRAINTS AND CONSUMER BEHAVIOR

Figure 5: Consumption growth per capita and efficiency unit in six OECD countries, 1950-2007.

Production grew at a lower rate than productivity (since \( \frac{c}{A} = \frac{C}{AN} \) declined) in European countries.

These figures on the other hand show that the theoretical predictions \( \dot{c}/c < g \) and \( \dot{l}/l < 0 \) from section 4.3 generally hold for European countries after 1950, but not for the United States. A similar conclusion emerges for the \( \dot{y}/y < g \) prediction.

Figure 6: Average labor hours supply per capita in six OECD countries, 1950-2007.

There exist theoretical explanations for some of these patterns. Prescott (2004), and the extension to subsistence consumption and more countries by Ohanian, Raffo and Rogerson (2008), argue that differences in labor supply variations displayed in figure 6 come mostly from changes in taxes\(^{24} \). Consumption patterns displayed in figure 5 have

\(^{24}\)There is a debate in the literature regarding calibration used in these models and Scandinavian outliers. Blanchard (2004) for instance remark that the elasticity of labor supply used by Prescott (2004) is large and consistent with macro but not micro evidence. Rogerson (2007) refines the analysis with home
not been considered but could also be explained by these institutional variations. Vandenbroucke (2009) numerically verifies another explanation for the decline in per capita labor supply: technological progress raised real wages and made leisure goods, which need time to be consumed, more affordable. The model presented in this paper is a special case, treating all goods as leisure goods. Proposition 2 thus represents an analytical proof of the finding of Vandenbroucke (2009) for this special case.

Other explanations have been provided for the cross-country variations in per capita labor supply but I am not aware of any study which considers cross-country variations in per capita consumption.

The consumption time model is partially consistent with this pair of macroeconomic evidence. Why the model is consistent with Europe and not the United States deserves further research. One possible explanation is the importance of consumption credit in the US, which could allow the consumption growth of non time-constrained low-income and middle-class households to compensate for the low consumption growth of time-constrained high-income households. Another possible explanation is the role of taxation. Personal income taxation indeed has been and is still more redistributive in the US than most countries in Europe, even if differences have narrowed. For instance, Wagstaff et al. (1999) calculate that the difference between the pre-tax and post-tax Gini coefficients in France was only -0.015 in 1989 while it was -0.038 in the US in 1987. Piketty and Saez (2007) calculate the pre-tax and post-tax income shares of the 10% richest households. The difference between the pre-tax and post-tax shares was only -1.1% in France in 2005, compared to -3.5% in the US in 2004. Differences in redistribution may be even larger if one take into account all taxes, including consumption taxes. Larger redistributive taxes in the US may mitigate more the effect of time constraints on rich households, as more of their income is redistributed to non-time constrained non-rich households.

8 Conclusion

Consuming takes time. Households in developing countries therefore face a trade-off as production technology improves: a high labor supply keeps production high but allows for little consumption of the additional production. Integrating the Becker (1965) home production and the Ramsey growth models, I showed that this trade-off results in a constantly decreasing labor supply and to a bound on consumption expenditures growth. Extending the model to heterogeneous agents, I also proved that rich save at a higher rate than poor households because coordination on the labor market makes them work more than desired and so they can not shift enough time to consume. The consumption time model also provides a simple explanation for the low observed rate of asset decumulation in old age and some other empirical evidence on consumer behavior.
Wealth is very concentrated in developed economies: 5% of the households hold more than 50% of assets in the United States. Even if in reality only few households are wealthy enough to lack time to consume, their decisions have a disproportionate impact on the aggregate capital stock. A good understanding of the savings behavior of the wealthy and how it impacts capital accumulation and output growth can thus help define macroeconomic policy. It is difficult to explain heterogeneous savings rates. Time need for consumption provides an additional complementary reason for the high saving rate of the wealthy.

To some degree the theory has policy implications in taxation. Time need for consumption is one reason for asset accumulation, but not the only one. For instance, a theory with finitely-lived agents and time need for consumption does not prevent bequest motives. It shows however that bequests can be, in part, accidental. Unlike the precautionary savings motive, the accidental part of bequests can be very large. This opens the door to alternative analysis of optimal capital and bequest taxation.

I finish with several additional research ideas. First, one can evaluate the quantitative relevance of the theory by predicting wealth distribution with numerical simulations and comparing the predictions with the empirical distribution, which is difficult to explain. The theory predicts that rich save at a higher rate, leading to wealth concentration. How much wealth concentration is an open question. Second, one could measure the empirical relevance of time constraints in actual consumption activities. The impossibility of running down assets by sole consumption is clear for multi-billionnaires but the question remains for less wealthy households. It would be interesting to identify time constraints and quantify their relevance from large household-level datasets on consumption, time use and wealth. Third, one could introduce taxation in the model and analyze the effect of taxes. Theoretical outcomes could be compared to macroeconomic evidence and potentially explain the differences between Europe and the United States in per capita output growth, relative to productivity growth. Finally, one could confirm empirically the theoretical prediction that labor supply declines with wealth on several margins. The theory is consistent along one of the extensive margin, retirement. One could use household-level data and treat causality carefully to test the prediction along the intensive margin, work hours, or additional extensive margins, such as participation.
Appendices
### A Proofs

More details are contained in Davoine (2011).

**Proof of proposition 1:** I first establish the dynamic system and analyze it, then show that \( \frac{\ddot{k}}{k} = k/Al \) increases over the long run, and finish by showing it implies that consumption expenditures will grow less than technological progress.

**Step 1.** Let us start with the dynamics system. It describes the evolution of labor supply \( l \), consumption \( c \) and capital \( \tilde{k} \), the latter two in per efficient labor unit term. To ease notation, I will not write conditions for \( l \). Looking at characteristics of optimal solutions for the dynamics system for \( \tilde{c} \) and \( \tilde{k} \) is sufficient to obtain the result. The capital law of motion is the same as in the Ramsey model: \( \dot{k} = f(\tilde{k}) - \tilde{c} - \left( \delta + g + \frac{\dot{\tilde{\epsilon}}}{\tilde{\epsilon}} \right) \tilde{k} \). However, the Euler equation is different (from lemma 1): \( \frac{\dot{\tilde{c}}}{\tilde{c}} = (r - \rho) / \varepsilon_{u_0}(c, l) + \frac{1}{\varepsilon_{u_0}(A_{\tilde{\ell}})} (f'(\tilde{k}) - \delta - \rho) - g + \left( \frac{\eta_0(A_{\tilde{\ell}})}{\varepsilon_{u_0}(A_{\tilde{\ell}})} - 1 \right) \frac{\dot{\tilde{l}}}{\tilde{\epsilon}} \). Writing \( Z \equiv (\tilde{c}, \tilde{k}, A, l) \), the dynamic system is thus

\[
\begin{pmatrix}
\dot{\tilde{c}} \\
\dot{\tilde{k}} \\
\end{pmatrix} = \mathcal{F}(Z) = \begin{pmatrix}
\mathcal{F}_1(Z) \\
\mathcal{F}_2(Z) \\
\end{pmatrix} = \begin{pmatrix}
\left[ \frac{1}{\varepsilon_{u_0}(A_{\tilde{\ell}})} (f'(\tilde{k}) - \delta - \rho) - g + \left( \frac{\eta_0(A_{\tilde{\ell}})}{\varepsilon_{u_0}(A_{\tilde{\ell}})} - 1 \right) \frac{\dot{\tilde{l}}}{\tilde{\epsilon}} \right] \tilde{c} \\
\frac{f'(\tilde{k}) - \delta - \rho}{\tilde{c}} - \left( \delta + g + \frac{\dot{\tilde{\epsilon}}}{\tilde{\epsilon}} \right) \tilde{k}
\end{pmatrix}
\]

On the associated \( \tilde{k} \times \tilde{c} \) phase diagramme, the \( \dot{\tilde{k}} = 0 \) and \( \dot{\tilde{c}} = 0 \) loci are moving over time (since \( A \) grows and \( l \) can change), so there are no steady-state constant solutions to the dynamic system (as there are in the Ramsey model). However, the intersection of these loci remain bounded, proving the existence and finiteness of solutions to the dynamic system.

**Step 2.** I show that \( \tilde{k} \) increases over the long run. Consider an increase of technology \( dA = A_1 - A_0 > 0 \) and the corresponding household labor supply choices \( dl = l_1 - l_0 \) (positive or negative). Let \( (\tilde{c}_i, \tilde{k}_i) \) be the solution to the system \( \dot{\tilde{c}}_i, \tilde{k}_i, A_i, l_i = \mathcal{F}(Z_i) = 0 \), which exist by step 1 \((i = 0, 1)\). I show that \( d\tilde{k} = \tilde{k}_1 - \tilde{k}_0 > 0 \) over the long run.

Making a Taylor expansion of \( \mathcal{F} \) around \( Z_0 \), we have \( 0 = \mathcal{F}(Z_1) \equiv \mathcal{F}(Z_0) + M(Z_1 - Z_0) = M(Z_1 - Z_0) \), where the matrix \( M \) takes values at the point \( Z_0 \) and is the gradient \( \nabla \mathcal{F}(Z_0) \). Writing explicitly this relation gives

\[
\begin{pmatrix}
0 \\
0
\end{pmatrix} \approx M \begin{pmatrix}
d\tilde{c} \\
d\tilde{k} \\
dA \\
dl
\end{pmatrix} = \begin{bmatrix}
\frac{\partial}{\partial \tilde{c}} \mathcal{F}_1 & \frac{\partial}{\partial \tilde{k}} \mathcal{F}_1 & \frac{\partial}{\partial A} \mathcal{F}_1 & \frac{\partial}{\partial l} \mathcal{F}_1 \\
\frac{\partial}{\partial \tilde{c}} \mathcal{F}_2 & \frac{\partial}{\partial \tilde{k}} \mathcal{F}_2 & \frac{\partial}{\partial A} \mathcal{F}_2 & \frac{\partial}{\partial l} \mathcal{F}_2
\end{bmatrix} \begin{pmatrix}
d\tilde{c} \\
d\tilde{k} \\
da \end{pmatrix}
\]

so, subtracting \( \frac{\partial}{\partial \tilde{c}} \mathcal{F}_1 / \frac{\partial}{\partial \tilde{c}} \mathcal{F}_2 \) times the second equation from the first equation and re-
arranging,
\[ d\tilde{k} \cong \frac{1}{\frac{\partial}{\partial k} \mathcal{F}_1 \frac{\partial}{\partial c} \mathcal{F}_2 - \frac{\partial}{\partial c} \mathcal{F}_1 \frac{\partial}{\partial k} \mathcal{F}_2} \times \left[ \left( \frac{\partial}{\partial A} \mathcal{F}_1 \frac{\partial}{\partial c} \mathcal{F}_2 - \frac{\partial}{\partial c} \mathcal{F}_1 \frac{\partial}{\partial A} \mathcal{F}_2 \right) dA + \left( \frac{\partial}{\partial l} \mathcal{F}_1 \frac{\partial}{\partial c} \mathcal{F}_2 - \frac{\partial}{\partial c} \mathcal{F}_1 \frac{\partial}{\partial l} \mathcal{F}_2 \right) dL \right] \]

One can show that, over the long run,

<table>
<thead>
<tr>
<th>Term</th>
<th>Long run value</th>
<th>Long run sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{\partial}{\partial k} \mathcal{F}_2 )</td>
<td>(</td>
<td>.</td>
</tr>
<tr>
<td>( \frac{\partial}{\partial k} \mathcal{F}_1 )</td>
<td>(</td>
<td>.</td>
</tr>
<tr>
<td>( \frac{\partial}{\partial A} \mathcal{F}_1 )</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>( \frac{\partial}{\partial A} \mathcal{F}_2 )</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>( \frac{\partial}{\partial l} \mathcal{F}_1 )</td>
<td>(</td>
<td>.</td>
</tr>
<tr>
<td>( \frac{\partial}{\partial l} \mathcal{F}_2 )</td>
<td>(</td>
<td>.</td>
</tr>
</tbody>
</table>

This follows from explicit calculations of the partial derivatives of \( \mathcal{F}_j \) and long run evolution of each of their respective terms. In particular, one can show that, as consumption expenditures \( c \) grow to infinity, \( \frac{\partial \epsilon_{u0}}{\partial c} \to 0 \), \( c \frac{\partial \epsilon_{u0}}{\partial c} \to 0 \), \( \frac{\partial \epsilon_{u0}}{\partial \tilde{l}_0} \to 0 \), \( \frac{\partial \eta_0}{\partial c} \to 0 \), \( c \frac{\partial \eta_0}{\partial c} \to 0 \), \( \epsilon_{u0} \to \frac{1}{\alpha} \), \( \eta_0 \to (\theta - \frac{1}{\alpha}) \frac{1}{1-\beta} \), \( \frac{\partial \eta_0}{\partial \tilde{l}_0} \to (\theta - \frac{1}{\alpha}) \left( \frac{1}{1-\beta} \right)^2 \) and \( \beta \to \sigma \theta \). Explicit differentiation of the labor schedule expression (5) and long run analysis deliver \( \frac{\partial}{\partial \tilde{l}_0} \left( \hat{l}_0/\tilde{l}_0 \right) \to 0 \), \( \frac{\partial}{\partial \tilde{l}_0} \left( \hat{l}_0/\tilde{l}_0 \right) \to -\frac{1}{\tilde{l}_0} \frac{1}{\sigma \theta} \Psi \), as well as another lengthy expression for \( \frac{\partial}{\partial k} \left( \hat{l}_0/\tilde{l}_0 \right) \).

For instance,
\[
\frac{\partial}{\partial A} \mathcal{F}_1 = \frac{\partial}{\partial \tilde{c}_0} \epsilon_{u0} (A_0 \tilde{l}_0 \tilde{c}_0) (\tilde{c}_0)^2 \tilde{l}_0 \left( f'(\tilde{c}_0) - \frac{\tilde{c}_0}{\epsilon_{u0} (A_0 \tilde{l}_0 \tilde{c}_0)} \right) + \frac{\tilde{c}_0}{\epsilon_{u0} (A_0 \tilde{l}_0 \tilde{c}_0)} \left( \tilde{c}_0 \frac{\tilde{l}_0}{\tilde{l}_0} \frac{\partial}{\partial c} \eta_0 (A_0 \tilde{l}_0 \tilde{c}_0) + \eta_0 (A_0 \tilde{l}_0 \tilde{c}_0) \frac{\partial}{\partial \tilde{l}_0} \left( \frac{\hat{l}_0}{\tilde{l}_0} \right) \right) - \tilde{c}_0 \frac{\partial}{\partial \tilde{l}_0} \left( \frac{\hat{l}_0}{\tilde{l}_0} \right) \\
\to 0 \times \text{bounded terms} \\
= 0.
\]

Hence,
\[
d\tilde{k} \to -\frac{1}{\frac{\partial}{\partial k} \mathcal{F}_1 \frac{\partial}{\partial c} \mathcal{F}_2} \left( \frac{\partial}{\partial \tilde{l}_0} \mathcal{F}_1 \frac{\partial}{\partial c} \mathcal{F}_2 \right) d\tilde{l} = -\frac{1}{\frac{\partial}{\partial \tilde{k}} \mathcal{F}_1 \frac{\partial}{\partial \tilde{l}_0} \mathcal{F}_2} \frac{\partial}{\partial \tilde{l}_0} \mathcal{F}_1 d\tilde{l}.
\]
and, using \( \text{sign} \left( \frac{\partial}{\partial l} F_1 \right) = \text{sign} \left( \frac{\dot{t}}{t} \right) = \text{sign} (dl) \),

\[
\text{sign} (d\tilde{k}) \rightarrow \text{sign} \left( -\frac{1}{\frac{\partial}{\partial \tilde{k}} F_1 \frac{\partial}{\partial l} F_1 dl} \right) = -\text{sign} \left( \frac{\partial}{\partial \tilde{k}} F_1 \right) \times \text{sign} \left( \frac{\partial}{\partial l} F_1 dl \right) = +
\]

so, over the long run, \( d\tilde{k} > 0 \).

**Step 3.** I conclude and show that \( \dot{c}/c < g \). Growth rate differentiation of the consumption expression in (5) and integration of the labor schedule expression gives

\[
\frac{\dot{c}}{c} = -\frac{l}{1-l} + \sigma \frac{\dot{w}}{w} = -\frac{1}{\beta} \Psi + \sigma \frac{\dot{w}}{w}.
\]

Over the long run, as consumption expenditures \( c \) grow, \( \varepsilon_u \rightarrow \frac{1}{\sigma} \) and \( \beta \rightarrow \sigma \theta \) so, plugging the \( \Psi \) expression and re-arranging,

\[
\frac{\dot{c}}{c} \rightarrow -\frac{1}{\sigma \theta} \left( \sigma \frac{\dot{w}}{w} + \frac{1}{\varepsilon_u} (r - \rho) \right) + \frac{\dot{w}}{w} = \frac{1}{\theta} (f'(\tilde{k}) - \delta - \rho) - \left( \frac{1}{\theta - \sigma} \right) \frac{\dot{w}}{w}.
\]

Since \( w = A(f(\tilde{k}) - f'(\tilde{k})\tilde{k}) \), we have \( \frac{\dot{w}}{w} = g - \frac{f''(\tilde{k})\tilde{k}}{f'(\tilde{k}) - f'(\tilde{k})\tilde{k}} \). Because \( \dot{k} > 0 \), \( f \) is concave and \( 1/\theta - \sigma > 0 \) (by assumption H1), \( \frac{\dot{w}}{w} \geq g > 0 \) and, over the long run,

\[
\frac{\dot{c}}{c} < \frac{1}{\theta} (f'(\tilde{k}) - \delta - \rho).
\]

Let \( k^* \) be the steady-state level of capital per efficient labor unit in the Ramsey model economy that shares the same attributes as the economy here, except there are no time constraints. Because \( \tilde{k} \) increases, we have \( \tilde{k} \geq k^* \) after some time, so \( f'(\tilde{k}) \leq f'(k^*) \) by concavity of \( f \). Using assumption H2, we thus have over the long run,

\[
\frac{\dot{c}}{c} < \frac{1}{\theta} (f'(\tilde{k}) - \delta - \rho) \\
\leq \frac{1}{\theta} (f'(k^*) - \delta - \rho) \\
= g.
\]

QED.

**Note:** the proof showed that \( \tilde{k} \) keeps increasing, so the interest rate \( f'(\tilde{k}) - \delta \) decreases. Could it be that it reaches a point where assumption H2 is violated? The answer is no, because there are also changes in \( \dot{l}/l \). Assume H2 is violated at some point. Then one can show that this implies \( \dot{l}/l > 0 \) the first time H2 is violated, when \( f'(\tilde{k}) - (\delta + g + \dot{l}/l) = 0 \). One can also show that \( \dot{l}/l \) is a continuous function and that \( \dot{l}/l < 0 \) as long as H2 holds, exhibiting a contradiction.
**Proof of proposition 3:** I first establish the dynamic system and analyze it, then show implications for the long run growth rate of consumption expenditures. Some steps are similar to the representative agent class, found in the proof of proposition 1.

**Step 1.** Let us start with the dynamics system and denote by $K = N^R a^R + N^P a^P$ the aggregate capital and $C = N^R c^R + N^P c^P$ aggregate consumption. As usual, tilde denotes per efficient labor unit values, so for instance $\tilde{k} = K/Al$ and $\tilde{c}^i = c^i/Al$. As in the proof of proposition 1, I do not write the conditions on labor supply $l$ in the dynamics system to ease notation. The capital law of motion is $\dot{K} = F(K, AL) - C - \delta K$ or $\dot{k} = f(\tilde{k}) - N^R \tilde{c}^R - N^P \tilde{c}^P - (\tilde{\delta} + g + \frac{L}{L}) \tilde{k}$. Using the Euler equation in lemma 3 and writing $Z \equiv (\tilde{c}^R, \tilde{c}^P, \tilde{k}, A, l)$, the dynamic system is then

$$
\begin{pmatrix}
\dot{\tilde{c}}^R \\
\dot{\tilde{c}}^P \\
\dot{\tilde{k}}
\end{pmatrix} = \mathcal{F}(Z) = 
\begin{pmatrix}
\frac{1}{\varepsilon_0(Alc^R, l)} (f'(\tilde{k}) - \delta - \rho) - g + \left( \frac{\eta_0(Alc^R, l)}{\varepsilon_0(Alc^R, l)} - 1 \right) \frac{i}{i} \tilde{c}^R \\
\frac{1}{\varepsilon_0(Alc^P, l)} (f'(\tilde{k}) - \delta - \rho) - g + \left( \frac{\eta_0(Alc^P, l)}{\varepsilon_0(Alc^P, l)} - 1 \right) \frac{i}{i} \tilde{c}^P \\
F(\tilde{k}) - N^R \tilde{c}^R - N^P \tilde{c}^P - \left( \tilde{\delta} + g + \frac{L}{L} \right) \tilde{k}
\end{pmatrix}.
$$

Similar arguments to the representative agent case show that there exist a unique solution to this system.

**Step 2.** Using an approach similar to the representative agent case and noting that $\frac{\partial}{\partial \tilde{c}} \mathcal{F}_1^R = 0$, the result of the Taylor expansion is

$$
0 \approx \frac{\partial}{\partial \tilde{c}^R} \mathcal{F}_1^R d\tilde{c}^R + \frac{\partial}{\partial \tilde{k}} \mathcal{F}_1^R d\tilde{k} + \frac{\partial}{\partial A} \mathcal{F}_1^R dA + \frac{\partial}{\partial l} \mathcal{F}_1^R dl,
$$

$$
0 \approx \frac{\partial}{\partial \tilde{c}^P} \mathcal{F}_1^P d\tilde{c}^P + \frac{\partial}{\partial \tilde{k}} \mathcal{F}_1^P d\tilde{k} + \frac{\partial}{\partial A} \mathcal{F}_1^P dA + \frac{\partial}{\partial l} \mathcal{F}_1^P dl,
$$

$$
0 \approx \frac{\partial}{\partial \tilde{c}^R} \mathcal{F}_2^R d\tilde{c}^R + \frac{\partial}{\partial \tilde{c}^P} \mathcal{F}_2^R d\tilde{c}^P + \frac{\partial}{\partial \tilde{k}} \mathcal{F}_2^R d\tilde{k} + \frac{\partial}{\partial A} \mathcal{F}_2^R dA + \frac{\partial}{\partial l} \mathcal{F}_2^R dl.
$$

To calculate each derivative element one note a difference with the 1 class case: in the 2 class case, the labor schedule at stage 3 (from definition 1) is taken as given by households, so their optimisation decisions will have no influence on the schedule. Hence, for any decision variable $x$, $\frac{\partial}{\partial x} \left( \frac{l}{l} \right) = 0$. Calculations and evaluations as in the representative case lead to $\frac{\partial}{\partial A} \mathcal{F}_1^R \rightarrow 0$, $\frac{\partial}{\partial A} \mathcal{F}_2^R = 0$, $\frac{\partial}{\partial \tilde{c}^P} \mathcal{F}_1^R \rightarrow 0$, $\frac{\partial}{\partial \tilde{k}} \mathcal{F}_2^R = 0$, $\frac{\partial}{\partial \tilde{c}^R} \mathcal{F}_1^R > 0$ and $\frac{\partial}{\partial \tilde{k}} \mathcal{F}_2^R < 0$. Subtracting $\frac{\partial}{\partial \tilde{k}} \mathcal{F}_1^R / \frac{\partial}{\partial \tilde{k}} \mathcal{F}_2^R$ times the third equation from the first equation and multiplying then by $\frac{\partial}{\partial \tilde{k}} \mathcal{F}_2^R$ gives

$$
d\tilde{c}^R \approx \frac{1}{\frac{\partial}{\partial \tilde{k}} \mathcal{F}_1^R \frac{\partial}{\partial \tilde{c}^P} \mathcal{F}_2} \left[ \left( \frac{\partial}{\partial \tilde{l}} \mathcal{F}_1^R \frac{\partial}{\partial \tilde{k}} \mathcal{F}_2^R \right) d\tilde{c}^P + \left( - \frac{\partial}{\partial \tilde{k}} \mathcal{F}_1^R \frac{\partial}{\partial \tilde{c}^P} \mathcal{F}_2^R \right) d\tilde{c}^P \right].
$$
and similarly for $d\tilde{c}^P$. Combining these two identities, one obtains an expression for $d\tilde{c}^R$ which does not depend on $d\tilde{c}^P$. Using the long run values of the derivative elements, I arrive at $d\tilde{c}^R \to 0$ and then $d\tilde{c}^P \to 0$. Hence, with $l/l < 0$ from lemma 3,

$$\frac{\dot{c}^i}{c^i} \to 0 \iff \frac{\dot{c}^i}{c^i} \to g + \frac{\dot{l}}{l} < g$$

QED.

**Proof of proposition 4:** I prove the result in three steps.

**Step 1.** $c^R/c^P < c^P/c^P$ over the long run. The Euler equations can be written $\dot{c}^i/c^i = G_3(c^i,l,r)$ with $G_1(c^i,l) = 1/\varepsilon_{u0}(c^i,l)$ and $G_2(c^i,l) = \eta_0(c^i,l)/\varepsilon_{u0}(c^i,l)$ and $G_3(c^i,l,r) = G_1(c^i,l) (r - \rho) + G_2(c^i,l) \dot{l}/l$. Because rich households have more labor income resources, they are in position to consume more than poor households at all time periods: $c^R > c^P$. I thus need to show that $\frac{\partial}{\partial c} G_3(c,l,r) < 0$, given schedules $l$ and $r$.

Using the functional forms for the household production function $h_0(c,l)$ and $\varepsilon_{u0}(c^i,l)$, calculations deliver the results. For instance, 

$$\frac{\partial}{\partial c} \varepsilon_{u0}(c,l) = \frac{\partial}{\partial c} \left[ \gamma \left( \frac{h_0(c)}{c} \right)^{1-\frac{\sigma}{\gamma}} \left( \varepsilon_{u0}(h_0,c) - \frac{1}{\sigma} \right) + \frac{1}{\sigma} \right] = \gamma \frac{1-\frac{\sigma}{\gamma}}{\sigma} \left( \frac{\theta}{c} - \frac{1}{\sigma} \right) \frac{1}{c} \left[ \gamma \left( \frac{h_0(c)}{c} \right)^{1-\frac{2\sigma}{\gamma}} - \left( \frac{h_0(c)}{c} \right)^{1-\frac{\sigma}{\gamma}} \right]$$

while, as $c$ grows, $\left( \frac{h_0(c)}{c} \right) \to 0$. When $c$ becomes large, both $\frac{\partial}{\partial c} G_1(c,l)$ and $\frac{\partial}{\partial c} G_2(c,l)$ are becoming small and tend to zero so a comparison of the speed at which they converge is needed. One obtains $\frac{\partial}{\partial c} G_1(c,l) \to M_1 \frac{1}{c} \left( \frac{h_0(c,l)}{c} \right)^{1-\frac{\sigma}{\gamma}}$ and $\frac{\partial}{\partial c} G_2(c,l) \to M_2 \frac{1}{c} \left( \frac{h_0(c,l)}{c} \right)^{1-\frac{\sigma}{\gamma}}$ where $M_1 < 0$ and $M_2$ are constant. Because of lemma 3, $\dot{l}/l < 0$, so $l$ declines over time and $\frac{\partial}{\partial c} G_2(c,l)$ declines faster than $\frac{\partial}{\partial c} G_1(c,l)$ over the long run. By assumption, $\left| \frac{\dot{l}}{l} \right| < (r - \rho)$ so $\frac{\partial}{\partial c} G_3(c,l,r) = \left( \frac{\partial}{\partial c} G_1(c,l) \right) (r - \rho) + \left( \frac{\partial}{\partial c} G_2(c,l) \right) \frac{1}{l} \to \left( \frac{\partial}{\partial c} G_1(c,l) \right) (r - \rho) < 0$.

**Step 2.** $y^R/y^P = y^P/y^P$ over the long run. By the capital law of motion,

$$\frac{\dot{a}^i}{a^i} = \frac{y^i - c^i}{a^i} = \frac{r d^i + w^i l - c^i}{a^i} = r + \frac{1}{a^i} (w^i l - c^i).$$

By proposition 3, $\dot{c}^i/c^i \to g + \dot{l}/l$ and by the constant skill premium assumption, $\dot{w}^i/w^i = \dot{w}/w$. Using similar steps as in the proof of proposition 1, one can show that $\dot{k} < 0$ over the long run. Since $w = A(f(k - f'(k))k$, we have that $\dot{w}/w = g - f''(k)kk/f'(k) - f'(k))$. By concavity of $f$, $\dot{w}/w < g$ so $w^i l$ grows at a lower rate than $g + \dot{l}/l$, so at a lower rate than $c^i$. Hence $(w^i l - c^i)$ decreases over time. Trivially the asset profile
\[ \dot{a}_i / a_i \geq 0. \]  
Since \( \dot{k} < 0 \), the interest rate \( r = f'(\bar{k}) - \delta \) increases over time. Summing up, \( \dot{a}_i \) increases over time, \( (w^l - c^l) \) decreases and \( r \) increases. At some point, \( \dot{a}_i / a_i = r + (w^l - c^l) / a_i \) will be defined by \( r \). But \( r \) is the same value for rich and poor households, so \( \dot{a}_R / a_R \rightarrow \dot{a}_P / a_P \).

Now pick one \( i \in \{ R, P \} \) and look at \( y_i = rd_i + w^l \). There are three cases to consider to obtain the long run growth of income \( y_i' / y_i \): either the terms \( rd_i \) and \( w^l \) grow at the same rate, or \( rd_i \) grows faster or \( rd_i \) grows slower. If terms \( rd_i \) and \( w^l \) grow at the same rate, their sum grows at the same rate, so \( \dot{y}_i / y_i = \dot{w}_i / w_i + l / l \). Since \( \dot{w}_i / w_i = \dot{w} / w \), \( y_R / y_R = y_P / y_P \). In the second case, ultimately the growth rate \( y_i' / y_i \) is defined by the growth rate of \( rd_i \). Since \( \dot{a}_R / a_R \rightarrow \dot{a}_P / a_P \), \( rd_R \) and \( rd_P \) grow at the same rate, so \( \dot{y}_R / y_R = \dot{y}_P / y_P \). In the third case, ultimately the growth rate \( y_i' / y_i \) is defined by the growth rate of \( wil \). Since \( \dot{w}_i / w_i = \dot{w} / w \), \( w_Rl \) grows at the same rate as \( w_Pl \), so \( \dot{y}_R / y_R = \dot{y}_P / y_P \). In all three cases, I have shown that

\[ \frac{y_R}{y_R} = \frac{y_P}{y_P}. \]

Implicitly, I have assumed that it is not possible to alternate between some of these three cases. This is indeed what happens, as calculations show that the difference between the growth rate of \( rd_i \) and that of \( w^l \) does not change sign.

**Step 3.** \( s^R > s^P \) over the long run. By definition \( s^i = \frac{y_i' - c_i}{y_i} = 1 - \frac{c_i}{y_i} \) so \( s^R > s^P \) if and only if \( c^R / y_R < c^P / y_P \). By the two first steps, \( c^R / c^R < c^P / c^P \) and \( \dot{y}_R / y_R = \dot{y}_P / y_P \). Hence

\[ \frac{\dot{c}_R}{c_R} \frac{y_R}{y_R} = \frac{\dot{c}_R}{c_R} - \frac{\dot{y}_R}{y_R} < \frac{\dot{c}_P}{c_P} - \frac{\dot{y}_P}{y_P} \]

so the growth rate of \( c^R / y_R \) is strictly lower than the growth rate of \( c^P / y_P \). Over the long run, we will thus have \( c^R / y_R < c^P / y_P \), so \( s^R > s^P \).

**QED.**

### B Life-cycle partial equilibrium version

Many empirical findings of consumer behavior are relevant in a life-cycle context. The theory developed in the main text is however cast in terms of dynasties. To allow a confrontation of this theory with empirical evidence, this appendix provides a life-cycle version of the consumption time model. This version will no longer hold in general equilibrium, as the interest rate will be taken as constant. Such a restriction is however without any consequences for the confrontation with the empirical findings considered here. Indeed, these always come from household level data and the empirical conclusion thus only hold in a partial equilibrium sense.

The adjustment consists in a basic life-cycle model in a partial equilibrium setting without uncertainty where households, past a certain level of wealth, are subject to a bound
on their consumption expenditures (positive) growth. Results of sections 4 and 5 are thus mirrored in this model if one assumes that the (partial equilibrium) interest rate is constant and the same as in a no-time constraints Ramsey model: consumption expenditures grow at a strictly lower rate than technological progress \( \dot{c}/c < g \); as in proposition 1), labor supply declines over time \( \dot{l}/l < 0 \); as in proposition 2) and rich save at a higher rate than poor from the same cohort \( s^R > s^P \); as in proposition 4).

Formally the life-cycle model takes place in the following partial equilibrium setting. The households have the same maximization problem as in the dynastic case, except that the maximization takes place from 0 to \( T < \infty \). This approach is similar to that followed in several theoretical other studies and yield the same Euler equation and first order conditions. The only differences on the consumption side are terminal conditions, but these were never used in the proofs in the dynastic setting case in sections 4 and 5. While terminal conditions in the dynastic and life-cycle cases are needed for a full characterization of the equilibrium, they are not needed for the results in this paper as these results are derived from some (but not all) properties of the equilibrium. In particular, relationships (5) still hold. The main difference on the production side is that we can no longer easily determine capital stock and interest rates endogenously, hence the partial equilibrium setting. It is sufficient to assume that the long run level of capital per efficient labor unit is not smaller than the one that would prevail in a world without time constraints to obtain results similar to propositions 1 and 2 \( \tilde{k} \geq k^* \) in the notation of the proofs in appendix A). For instance, one can simply assume that they are equal. Then the results remain (given the assumption \( \tilde{k} \geq k^* \), it is sufficient to perform a small adjustment of step 3 of the proof for proposition 1, namely \( \tilde{k} = 0 \) so \( \dot{w}/w = g > 0 \). A constant value \( \tilde{k} = 0 \) is sufficient to prove results which lead to proposition 4.

Note that I could add uncertainty or other realistic features in the model but I refrain from doing so to isolate the effect of time constraints. Also note that the results in the life-cycle version of the model are proved under partial equilibrium and could also hold in general equilibrium. Proofs are however needed for the general equilibrium case.

Finally, note that intuition for the consumer behavior in this life-cycle partial equilibrium setting is the same as the intuition in the dynastic general equilibrium version of the main text. Intuitively, the same mechanisms operate but in a more acute way: when time lacks for consumption, finite life makes the lack of time even more stringent than in a dynastic setting.
Bibliography


[77] Li, Geng (2009), Transaction costs and consumption, *Journal of Economic Dynamics and Control*, vol. 33(6), 1263-1277.


Chapter 4

Intertemporal trade, Sovereign risk premiums and Eurobonds

Thomas Davoine

Vienna, October 2012, updated May, September 2013

Abstract

I extend a basic overlapping generations model to analyze the impact of cross-country policies in a common currency area, including intertemporal trade, endogenous labor supply and incomplete debt contracts. I compare the outcome of standard domestic policy reforms to transfers between governments and the use of a common bond. Loans to governments are not fully protected against defaults and thus come with an interest rate premium, which I derive assuming uncorrelated shocks and abstracting from moral hazard. Common bonds help countries with low public debt sustainability finance their debt at a lower cost, change households decisions in the beneficiary countries and thus the equilibrium interest rate and trade balance patterns, leading to spillover effects which may also benefit countries who agree to pay a higher debt premium. I analytically show that the public debt premium is convex decreasing in the level of debt sustainability under certain conditions. Numerical simulations show that common bonds help improve debt sustainability but at a welfare cost over the long run, as private investments lead to large trade re-balancing effects. Moderate fiscal consolidation improve both sustainability and welfare, thanks to a multiplier effect


Keywords: endogenous public debt premium, OLG, intertemporal trade, Eurobonds, fiscal consolidation

---

1University of St.Gallen and Institute for Advanced Studies, Vienna. Contact: thomas.davoine@unisg.ch. Comments by seminar participants at the University of St.Gallen and the 2012 LMU Public Finance seminar in Munich are gratefully acknowledged.
1 Introduction

Historically, most of policy interventions has been taking place at the country level. As trade develops however, the economic futures of countries have been increasingly linked together, triggering more discussions about cross-border policies that go beyond trade and product market regulation. Member countries of the European Union (EU) for instance have recently been debating the creation of joint government bonds (Eurobonds). The aim of this paper is to build an analytical framework to be able to evaluate cross-border policies within common currency areas. Loans to governments are not fully protected against defaults and thus come with an interest rate premium, which decrease with debt sustainability. Eurobonds help countries with low debt sustainability finance their debt at a cheaper rate. Households of beneficiary countries may then change labor supply and savings decisions, the equilibrium interest rate and trade balance patterns may adjust, leading to spillover effects which may also benefit countries which in the first place agreed to pay a higher debt premium. Under simplifying assumptions, I derive the public debt premium and show analytically that the premium can be convex. Numerical simulations show that Eurobonds help improve debt sustainability but at a long run welfare cost.

I will compare cross-country policies to domestic policies which can also improve debt sustainability: cross-country transfers, Eurobonds and fiscal consolidation reforms. Recent empirical analysis has shown that Eurozone cross-border transfers (the so-called structural funds) have had a positive impact on the rate of growth of recipient countries, even if moderate (e.g. Becker, Egger and von Ehrlich, 2010). The use of Eurobonds to resolve the current public debt crisis in some European countries is a topic of debate. Muellbauer (2011) for instance advocates the introduction of conditional Eurobonds with an incentive mechanism for reforms, addressing moral hazard concerns.

To analyze cross-border policies, I build a model which is an extension of a standard Blanchard (1985) overlapping generations model with endogenous labor supply and intertemporal trade.

Intertemporal trade takes place between many countries, following the two-country Buiter (1981) approach with a universal composite good, common in intertemporal approaches to current account analysis (Obstfeld and Rogoff, 1995). The choice of a single composite good is made to isolate the effect of public policy differences from other trade effects, such as specialization. One contribution to the literature is the provision of a quantitative criteria to identify winners and losers from intertemporal trade. Thanks to mobile capital, households in the low autarky interest rate have access to better rewards in the other countries, the resulting trade flows modify capital accumulation patterns, the production levels may change and lead to either a welfare gain or a welfare loss. With only two countries the criteria can be applied ex-ante. Unlike Buiter (1981), intertemporal trade does not only result from differences in time preferences across countries: the interest rate wedge which gives rise to trade flows can also result from differences in
production or domestic policy.
The extension addresses incomplete debt contracts. In case of disputes, loans to government have less protection than loans to domestic households. Assuming otherwise perfect capital markets and uncorrelated shocks to government expenditures, I derive the public debt premium from microfoundations. Following Hamilton and Flavin (1986) and consistent with Frenkel and Razin (1986), I define the sustainable level of debt as the expected present value of future fiscal surpluses. As in Uribe (2006), governments will default if shocks push debt beyond the sustainable level. I analytically show that the public debt premium decreases with the sustainable debt level, sometimes in a convex manner.

Consistent with generational accounting approaches, the cost of public debt depends not only on the past (accumulated debt) but also on the future (expected future surpluses). The resulting debt premium is more likely to explain why in reality countries with similar debt-to-GDP ratios may pay different interest rates on their debt, a contradiction for methods which only take past information into account.

The approach is also consistent with the sovereign debt literature but uses an endogenous characterization of public debt premiums. In general, premiums on the sovereign default risk are quantified via reduced-form specifications, either directly (as in Schmitt-Grohé and Uribe, 2003; Corsetti, Kuester, Meier and Mueller, 2013) or via output costs of defaults (as in Arellano, 2008).

The microfoundation basis that I use is the default mechanism from Uribe (2006), extended in ways similar to Juessen, Linnemann and Schabert (2010), Grossmann (2011) and Kohler (2012). In contrast to these three studies however, I assume that perfect competition holds for all firms, including financial intermediaries. As a result, there are no multiple equilibrium nor self-fulfilling properties. One contribution of this paper is an explicit public debt premium function along with analytical convexity properties.

My model is closest to the one from Juessen, Linnemann and Schabert (2010). In their model, defaults occur when shocks push government debt beyond the net present value of future expected fiscal surpluses. Because default expectations influence the price of debt and the discounting factor, multiple equilibrium and self-fulfilling expectations are possible in their model: expectations of higher default increase the price of debt, which increase the likelihood of defaults. My model is a variation of their model, with three essential differences. First, I use a multi-country model to investigate cross-country policy effects. Second, I use an overlapping-generations structure, rather than infinitely-lived agents. Persson (1985) and Frenkel and Razin (1986) have indeed shown that intertemporal terms-of-trade effects play an important role in international spillovers of fiscal policy: both interest rate and trade balance variations matter. Yet infinitely-lived agents settings constrain interest rate variations. While this may be appropriate in a closed economy, it is no longer the case for multi-country analysis. Third, I modify the debt premium analysis: in particular, I assume uncorrelated shocks and perfect competition in the debt market, leaving no room for multiple equilibrium.
There are only a few quantitative studies of intertemporal trade between open economies allowing to isolate the effect of policy differences. The most relevant are Frenkel, Razin and Symansky (1991), Jäger and Keuschnigg (1991), Mendoza and Tesar (1998) and Hayakawa and Zak (2002). All consider spillover effects of tax reforms. However, none uses endogenous costs of public debt.

As in Juessen, Linnemann and Schabert (2010) and Corsetti, Kuester, Meier and Mueller (2013), the convexity of the premium function plays an important role: public debt improvements at low debt sustainability levels have a larger impact, leading to a variety of spillover and intertemporal terms-of-trade effects. For numerical simulations I derive a convex debt premium function from the calibration and specification of interest rate premiums used by Schmitt-Grohé and Uribe (2003). Simulations show that cross-country transfers only lead to long-run welfare improvements when they are small, a result consistent with the transfer paradox literature (Galor and Polemarchakis, 1987). Eurobonds achieve their goal of improving the sustainability of public debt but at a long-run welfare cost: increasing public debt crowds-out productive capital at home, creates an interest rate wedge between countries which attracts more investments from foreign countries, a trade re-balancing effect which builds foreign debt and depletes household wealth, welfare and consumption over the long run. On the other hand, thanks to a multiplier effect that they provide, moderate fiscal consolidation measures achieve an improvement in both debt sustainability and welfare. Simulations identify redistribution effects of Eurobonds, from future domestic generations to current domestic generations and future foreign generations, although they do not separate efficiency losses from redistribution. Simulations also illustrate the cross-country insurance potential of Eurobonds.

The next section presents the model. Section 3 provides theoretical results. Section 4 presents numerical simulation results. The last section concludes.

2 Model

The basis of the model is an overlapping generations model with stochastic death (Blanchard, 1985) and endogenous labor supply. Its intertemporal trade extension assumes that $M$ countries produce the same composite good (Buiter, 1981). Perfect mobility of savings and no arbitrage implies that households in different countries face the same interest rate. Governments uses labor and consumption taxes to finance their own consumption and can accumulate debt but have to pay a premium to cover the sovereign debt default risk, resulting from the incompleteness of public debt contracts\(^2\).

\[^2\text{For convenience, I present a simple version of the model. Numerical simulations will use a more complete model. Theoretical results hold in both versions.}\]
2. MODEL

**Households** Population is of constant size \( N_i \) in each country \( i \in \{1, \ldots, M\} \). Households are identical but are born at different times and die at different ages. They discount time with factor \( \beta \) and face a constant probability of dying \( 1 - \gamma \). As in Blanchard (1985), I assume that there is an actuarially fair reverse-life insurance so households have the budget constraint

\[
\gamma A_{v,t+1} = R_{t+1} (A^i_{v,t} + (1 - \tau^i_v) w^i_{v,t} l^i_{v,t} - (1 + \tau^i_c) c^i_{v,t}),
\]

where \( A^i_{v,t} \) are the assets at time \( t \) of a household born at time \( v \), \( c^i_{v,t} \) is consumption of the household, \( l^i_{v,t} \) is labor supply of the household, \( R_{t+1} = 1 + r_{t+1} \) is the gross interest rate, \( \tau^i_v \) is the tax rate on labor income, \( \tau^i_c \) is the tax rate on consumption and \( w^i_{t} \) is the wage rate.

Households choose how many work hours to supply. They have a preference for leisure, represented by a constant-elasticity labor disutility function

\[
\phi(l^i_{v,t}) = \phi_0^{-1/\epsilon} (l^i_{v,t})^{1+1/\epsilon} / 1 + 1/\epsilon
\]

expressed in goods-equivalent terms. Preferences are defined over consumption net of labor disutility, or net consumption in short, \( Q^i_{v,t} = c^i_{v,t} - \phi(l^i_{v,t}) \). Expected lifetime utility is given recursively by \( U^i_{v,t} = u(Q^i_{v,t}) + \beta \gamma U^i_{v,t+1} \), where the instantaneous utility has as an isoelastic form \( u(Q) = (Q^{(\sigma-1)/\sigma} - 1)/(\sigma - 1) \). Households thus make consumption and savings decisions in order to solve the following (recursive) maximization problem:

\[
U(A^i_{v,t}) = \max_{l^i_{v,t}, c^i_{v,t}} \left\{ u(Q^i_{v,t}) + \beta \gamma U(A^i_{v,t+1}) \right\}
\]

s.t. \( \gamma A_{v,t+1} = R_{t+1} (A^i_{v,t} + (1 - \tau^i_v) w^i_{v,t} l^i_{v,t} - (1 + \tau^i_c) c^i_{v,t}) \)

\( Q^i_{v,t} = c^i_{v,t} - \phi(l^i_{v,t}) \)

**Firms** Firms are owned by households. Production in each country uses two factors with exogenous productivity \( A^i_o \) and is represented by a Cobb-Douglas function

\[
Y^i_t = F^i(K^i_t, L^i_t) = A^i_o (K^i_t)^\alpha (L^i_t)^{1-\alpha}.
\]

In competitive markets, factors are paid their marginal products \( w^i_t = \partial F^i / \partial L^i_t = (1 - \alpha)Y^i_t / L^i_t \) and \( r^i_t = \partial F^i / \partial K^i_t - \delta = \alpha Y^i_t / K^i_t - \delta \), recalling that interest rates are equal across countries and noting \( \delta \) for the capital depreciation rate. Labor markets are perfect so labor demand equals labor supply, \( L^i_t = \sum_v N^i_{v,t} l^i_{v,t} \), where \( N^i_{v,t} \) is the number of

\(^3\)In the notation, superscripts \( i \in \{1, \ldots, M\} \) indicate the country. Except the interest rate, every parameter or variable can differ between countries. For ease of reading, I only indicate the country for the most important parameters.
households alive at time \( t \) and born at time \( v \). In contrast with part of the literature which extended Buiter (1981), I allow production technologies to differ across countries and therefore can not claim that wages are also equalized across countries. Capital follows the low of motion

\[
K_{t+1}^i = (1 - \delta)K_t^i + I_t^i,
\]

where \( I_t^i \) denotes investment. Firms take wages \( w_t^i \) and labor supply \( L_t^i \) as given, make profits \( \Pi_t^i = F^i(K_t^i, L_t^i) - w_t^i L_t^i \) and distribute dividends \( \Pi_t^i (K_t^i) - I_t^i \). The firm value \( V_t^i \) is the discounted sum of accumulated dividends, which one can represent in dynamic form by\(^4\):

\[
V_t^i(K_t^i) = \max_{I_t^i} \Pi_t^i (K_t^i) - I_t^i + \frac{V_t^i(K_{t+1}^i)}{R_{t+1}}.
\]

**Governments**  In the model, policy is exogenously defined. At country level, government raises labor income tax at rate \( \tau_t^i w \) and consumption tax at rate \( \tau_t^i c \) in order to finance its own consumption \( CG_t^i \). At the cross-country level, governments can agree on cross-subsidies or common bonds (Eurobonds). I denote by \( Z_t^i \) the total net subsidy that country \( i \) decides to make to other countries\(^5\). Subsidies are made from government to government so only enter the government budget constraints. Governments can run budget deficits and finance them with debt, paying a premium \( P_t^i \) on top of the interest rate \( R_{t+1} \) when there is no Eurobond. The premium \( P_t^i \) covers the sovereign debt default risk and will be derived below. I denote by \( \mathcal{E}_t^E = 1 \) the presence of a Eurobond and \( \mathcal{E}_t^E = 0 \) its absence. In case of a Eurobond, governments finance their debt at the same interest rate \( R_t + \bar{\pi} \) for a premium that lies somewhere between all premiums\(^6\), that is \( \min\{P_t^i\} \leq \bar{\pi} \leq \max\{P_t^i\} \). Debt \( DG_t^i \) and budget surplus \( \mathcal{J}_t^i \) are then given by:

\[
DG_{t+1}^i = \left( R_{t+1} + \bar{\pi} \mathcal{E}_t^E + P_t^i (1 - \mathcal{E}_t^E) \right) \times (DG_t^i - \mathcal{J}_t^i)
\]

\[
\mathcal{J}_t^i = \tau_t^c C_t^i + \tau_t^w w_t^i L_t^i - CG_t^i - Z_t^i.
\]

Following Uribe (2006), Juessen, Linnemann and Schabert (2010) and Corsetti, Kuester, Meier and Mueller (2013), defaults occur mechanically, so there is no need to model a government motive and explicit decisions. Two reasons motivate this approach. First, government keep taxes constant and government expenditures will be submitted to random shocks so defaults will occur automatically once a threshold on debt has been met. I will be more specific below. Second, Yeyati and Panizza (2011) provide empirical evidence indicating that output costs of sovereign defaults take place before defaults.

---

\(^4\)In the complete setting for numerical simulations, I add capital adjustment costs.

\(^5\)When needed, I will denote by \( Z_t^{ij} \) the subsidy made by country \( i \) to country \( j \), with \( Z_t^i = \sum_{j \neq i} Z_t^{ij} \).

When \( Z_t^{ij} > 0 \), country \( i \) makes a transfer and when \( Z_t^{ij} < 0 \), the subsidy is received.

\(^6\)In the quantitative analysis, \( \bar{\pi} \) will be the weighted sum of premiums \( P_t^i \), using debts as weights. For the theoretical analysis, I will assume that governments ignore changes in \( \bar{\pi} \) and take it as constant.
Because the penalty takes place before default decisions, I assume that governments and countries suffer from high debt levels in a continuous fashion and abstract from explicit government default decisions. This approach however does not allow for an analysis of moral hazard, which can arise for instance if highly indebted countries promise tax reforms in exchange for the immediate creation of a Eurobond.

**Incomplete public debt contracts**  In practice, we observe large differences on public debt interest rates, even with the same currency (see figure 1). These differences can be justified by a risk premium, some government defaulting or delaying debt repayments. Investors lending to government are indeed not protected against government defaults. Since they are not perfectly enforceable, debt contracts are incomplete, leading investors to charge a risk premium.

![Figure 1: Interest rates on selected government bonds (maturity 10 years)](image)

My approach is similar to that of Juessen, Linnemann and Schabert (2010), Grossmann (2011) and Kohler (2012). I present my approach first and return to differences later.

I assume that governments expenditures $CG_i$ are random, independently and identically distributed across countries (normalized by the economic size of the country) and write $\phi(E[CG_i]/Y_i, \bar{s}^2)$ for the cumulative distribution function representing the random process $CG_i/Y_i$ with variance $\bar{s}^2$.

Expenditure shocks capture a variety of real-life events affecting fiscal balances, such as health expenditure shocks (e.g. following a climate shock), social insurance shocks (e.g. extension of unemployment benefits under strong negative business cycle shock) or demographic shocks (e.g. baby-boom generation retirement under pay-as-you-go pensions).

---

7The sovereign debt literature makes a generic usage of the words “incomplete debt contracts”, without referring specifically to microeconomic contract theory, and I follow this practice (see for instance Tomz and Wright, 2013).
Because of uncertainty in government expenditures, the fiscal surplus can vary and the path of government debt $DG_{it}^j$ is uncertain. Debt could grow to very large levels but I follow Uribe (2006) in assuming that investors will simply not lend to governments for debt levels which are \textit{unsustainable} over the long run, because they see it as a guarantee for default. Investors define the sustainable debt level and charge a premium depending on this level. Investors take observed current fiscal policy as granted, because they do not believe in government promises: investors will re-adjust their premiums after a fiscal reform decisions but not before. I denote by $DG_{it}^{s,j}$ the sustainable debt level and follow the so-called \textit{present-value borrowing constraint for government debt} from Hamilton and Flavin (1986) for its definition: $DG_{it}^{s,j}$ is the debt level such that the present value of government expenditures equals the expected present value of government revenue, taking into account the total cost of the debt $R_{t+1} + P_t^i$. In other words, $DG_{it}^{s,j}$ is equal to the expected present value\textsuperscript{8,9} of fiscal surpluses $S_{it}$:

\begin{equation}
DG_{it}^{s,j} = NPV(E[S_{it}]) \nonumber \\
= E \left[ \sum_{u=t}^{\infty} \left( \prod_{v=t}^{u-1} \frac{1}{R_{v+1} + P_v} \right) S_{u} \right] \nonumber \\
= \sum_{u=t}^{\infty} \left( \prod_{v=t}^{u-1} \frac{1}{R_{v+1} + P_v} \right) \left( \tau_e^i C_u^i + \tau_w^i W_u^i L_u^i - (E[C_{G_{u}}^i] + Z_u^i) \right) .
\end{equation}

Because governments can not finance any debt level on top of the sustainable level defined by investors but can choose to borrow less, it always holds that

$$DG_{it}^j \leq DG_{it}^{s,j}.$$ 

I will refer to $DG_{it}^{s,j}$ as the \textit{public debt capacity}, the maximum level beyond which investors will not lend. When the size of the shock on $CG_{it}^j$ is large and brings the debt level $DG_{it}^j$ above the debt capacity $DG_{it}^{s,j}$ there is a mechanical default by government, as in Uribe (2006)\textsuperscript{10}.

The approach I use is a variation of Juessen, Linnemann and Schabert (2010). I borrow the fact that government bonds are non-state contingents, the mechanical default mechanism and the role of future expected fiscal surpluses in defaults. I also consider stationary steady-state outcomes. On the other hand, I always consider uncorrelated shocks, while they do it for analytical results but not for numerical results. Investors take the current tax rate for granted, rather than considering hypothetical and revenue-

\textsuperscript{8}Aggregate consumption in country $i$ at time $t$ is $C_t^i = \sum_{v=t}^{\infty} N_{v,t}^i c_{v,t}^i$, and similarly for other variables.

\textsuperscript{9}Shocks to $CG_{it}^j$ propagate to the actual debt level $DG_{it}^j$ and other economic variables, including $C_t^j$. One should thus write $E[C_t^j]$. For a lighter notation, I only write $C_t^j$ and similarly for other variables to refer to their stationary steady-state value.

\textsuperscript{10}Under certain conditions, I will show that $E[DG_{it}^j] = DG_{it}^{s,j}$, so the expected debt level is the discounted value of future expected surpluses, consistent with the public debt definition of Frenkel and Razin (1986).
maximizing levels in calculating future expected fiscal surpluses. Finally, I will assume perfect capital markets in my analysis, which I will show leads to single equilibrium and prevents self-fulfilling prophecies.\footnote{On top of public debt modeling, there are other differences: I consider overlapping-generations, not infinitely-lived agents; production factors include capital, not just labor; and my model is multi-country.}

The approach is also similar to Grossmann (2011) and Kohler (2012). Default mechanisms are identical, including the role of future expected fiscal surpluses. However, I assume perfect capital markets, which will rule out multiple equilibrium and allow for an explicit, rather than implicit, characterization of the public debt premium.\footnote{As I will elaborate below, trade in my model takes place for the same composite good, which I use as numeraire. Exchange-rate variations then play no role. The interest rate parity condition from Grossman (2011) will then correspond to my condition of equalization of interest rates across countries.}

I will discuss this assumption later.

In section 3, I will derive the premium $P^i_t$ from the probability of default $P(DG^i_t > DG^{s,i}_t)$ and will show that, assuming otherwise perfect capital markets and other realistic conditions, the premium will only depend on the (normalized) debt capacity level and will be nothing but an insurance premium\footnote{The present-value borrowing constraint for government debt used to define debt sustainability is different from \textit{natural borrowing limits}, which refer to worst case realizations as opposed to expectations. Budget constraints do not hold in models with defaults. The premium function $\pi(\cdot)$, which acts as an insurance premium, however, captures the cost of defaults. Then the budget constraint holds.} which can be represented by a function $\pi(\cdot)$ such that

$$P^i_t = \pi \left( \frac{DG^{s,i}_t}{Y^i_t} \right).$$

### Intertemporal trade
Following the two-country approach from Buiter (1981), I assume that all $M \geq 2$ countries produce the same composite good and can trade either this good or capital. Labor is immobile. There is no trade with the rest of the world.\footnote{The Buiter (1981) approach is standard in intertemporal approaches to current account analysis (section 3, Obstfeld and Rogoff, 1995) and already used for $M \geq 2$ countries with overlapping generations (for instance, Börsch-Supan, Ludwig and Winter, 2006).}

The $M$ countries engage in trade and keep trade balances, defined by

$$TB^i_t = Y^i_t - C^i_t - CG^i_t - Z^i_t - I^i_t.$$ 

The part of the output that is not consumed, invested or transferred to other governments is exported (respectively imported, if output is low). The net foreign asset position $DF^i_t$ enters total household wealth and follows the law of motion\footnote{In the complete setting for numerical simulations, I add trade costs.}

$$DF^i_{t+1} = R_{t+1} \left( DF^i_t + TB^i_t \right).$$
dition is:

\[ A_i^t = DF_i^t + DG_i^t + V_i^t. \]

When there is no trade, interest rates differ in each country and adjust so that the goods market in each country clears (equivalently, \( TB_i^t = 0 \)). When there is trade, goods markets at the country level do not clear (\( TB_i^t \neq 0 \)). Because there is no trade with the rest of the world, the overall goods market clears over the \( M \) countries, as the unique interest rate adjusts:

\[ \sum_{i=1,...,M} TB_i^t - TC_i^t = 0. \] (4)

**Remarks** It is possible to express the public debt premium in terms of the implicit debt level used in generational accounting (Auerbach, Gokhale and Kotlikoff, 1992). Both the model presented here and generational accounting take future commitments into account when defining public debt levels (for technical details, see Davoine, 2013). In reality, there are several reasons for countries to trade. Using the composite goods approach of Buiter (1981) and the intertemporal current account literature, the model captures differences in preferences, technology and policy but does not capture specialization. Because I want to compare the effect of different domestic and cross-country policies, it is sufficient.

### 3 Theoretical results

The two main analytical contributions of this paper are the characterization of the public debt premium and the characterization of benefits from intertemporal trade. Before deriving them, I consider optimal household behavior and obtain results which are not new, but helpful for the analysis\(^{16}\).

#### 3.1 Household decisions

I characterize optimal decisions by households and their impact at the aggregate level. The interest rate is always taken as given by households. The first order conditions of the household maximization problem (2) are

\[ u'(Q_{i,t}) = (1 + \tau_i^c)\beta R_{t+1}U'(A_{i,t+1}) \]

\[ u'(Q_{i,t}) \phi'(l_{i,t}) = \beta R_{t+1}U'(A_{i,t+1}) (1 - \tau_i^w)w_i. \]

Using the functional forms for \( u \) and for \( \phi \) and the envelope theorem, one obtains the following characterization of household decisions:

\(^{16}\)Other theoretical results can be obtained in a small open economy setting, which may be interesting for discussions focused on trade balance effects. In particular, one can not improve trade balance and welfare at the same time by introducing Eurobonds.
Lemma 1. (Household decisions) optimal net consumption, labor supply and goods consumption evolve according to:

\[ Q^{i}_{t+1} = (\beta R_{t+1})^{\sigma} Q^{i}_{t}, \quad l^{i}_{t+1} = l^{i}_{t} \equiv \varphi_{0} \left[ \frac{1 - \tau^{w}_{i} w^{i}_{t}}{1 + \tau^{c}_{i}} \right]^{\epsilon}, \quad c^{i}_{t+1} = Q^{i}_{t+1} + \varphi(l^{i}_{t}). \]

Note that the result includes the Euler equation and that households provide the same labor \( l^{i}_{t} \) whatever their age. As expected, higher wages \( w^{i}_{t} \), lower labor income taxes \( \tau^{w}_{i} \) and lower consumption taxes \( \tau^{c}_{i} \) all lead to higher labor supply, ceteris paribus.

Using the Cobb-Douglas functional form for production and the fact that factors are paid their marginal product, algebra allows to derive aggregate labor supply variations \( L^{i}_{t} = N^{i} l^{i}_{t} \) from individual labor supply decisions. One obtains:

Lemma 2. (Aggregate labor supply): given capital stock \( K^{i}_{t} \), households will in aggregate provide the following labor supply:

\[ L^{i}_{t} = \left[ N^{i} \varphi_{0} \left( \frac{1 - \tau^{w}_{i} A^{i}_{0}(1 - \alpha)}{1 + \tau^{c}_{i}} K^{i}_{t} \right)^{\alpha} \right]^{1/(1+\alpha \epsilon)}. \]

Deriving aggregate households consumption from individual decisions is more involved, but purely technical. Derivations do not deliver any particular insights and are left for the technical appendix Davoine (2013). Households decisions lead to the following evolution of aggregate consumption:

Lemma 3. (Aggregate consumption): Aggregate net consumption in country \( i \) evolves according to:

\[ Q^{i}_{t} = \frac{1}{(1 + \tau^{c}_{i})\Omega^{i}_{t}} (A^{i}_{t} + H^{i}_{t}). \]

Moreover, the following budget constraint and relationships hold:

\[ A^{i}_{t+1} = R_{t+1} (A^{i}_{t} + (1 - \tau^{w}_{i}) w^{i}_{t} L^{i}_{t} - (1 + \tau^{c}_{i}) C^{i}_{t}) \]
\[ \Omega^{i}_{t} = 1 + \gamma \beta^{\sigma} R^{-1}_{t+1} \Omega^{i}_{t+1} \]
\[ H^{i}_{t} = (1 - \tau^{w}_{i}) w^{i}_{t} L^{i}_{t} - (1 + \tau^{c}_{i}) N^{i} \varphi(L^{i}_{t}/N^{i}) + \frac{H^{i}_{t+1}}{R^{t+1}} \]
\[ C^{i}_{t} = Q^{i}_{t} + N^{i} \varphi(L^{i}_{t}/N^{i}). \]

Interpretation of lemma 3 is useful for the analysis of quantitative results. The last relationship of the lemma allows conversion of consumption \( C^{i}_{t} \) into net consumption \( Q^{i}_{t} \), and vice-versa. Long run welfare effects of policy reforms can be evaluated by their impact\(^{17}\) on \( Q^{i}_{t} \). In the reminder of the text, I will thus use interchangeably the terms net

\(^{17}Q\) is an appropriate measure of long-run welfare impacts for the following reasons. In a steady-
consumption and welfare.

$A_t^i$ represent financial assets owned by the households, the sum of firm value, public debt and foreign assets, $A_t^i = V_t^i + DG_t^i + DF_t^i$. The aggregate $H_t^i$ represents human wealth, the net present value of future labor income streams, net of taxes and disutility of effort. The sum $A_t^i + H_t^i$ represents total private wealth and $\Omega_t^i$ the inverse of the marginal propensity to consume, net of consumption taxes.

Changes in household wealth are thus a convenient characterization of long run welfare impacts, taking into account intertemporal adjustments. When $\Omega_t^i$ is constant (e.g. with a logarithmic utility, $\sigma = 1$), a reform has a positive impact on welfare if and only if it has a positive impact on total private wealth. For instance, decreases in labor income taxes $\tau_t^w$ push labor supply $L_t^i = N_t^i l_t^i$ up as well as human wealth $H_t^i$, ceteris paribus. Higher households total wealth would push welfare $Q_t^i$ and consumption $C_t^i = Q_t^i + N_t^i \varphi(L_t^i/N_t^i)$ up.

While lemmas 1 and 2 are direct applications of first order conditions and may be familiar, lemma 3 is less frequent. It is however not a novelty: theoretical analysis by Frenkel and Razin (1986) and numerical simulations by Jäger and Keuschnigg (1991) also decompose consumption choices as the product of a marginal propensity to consume with total household wealth.

### 3.2 Endogenous public debt premium

This section establishes existence of an endogenous risk premium function for public debt and provides some of its properties, when there is no Eurobonds. I start by expressing the probability of default in terms of the distribution of government expenditure shocks and then move to the expression of the debt premium function, before providing characteristics of this function. I also make some simple investigations of public debt changes caused by policy reforms, in order to build an understanding of the mechanisms at work. For a lighter notation, I drop the country index $i$.

I now express the probability of government default in terms of the distribution of government expenditure shocks, considering the case of constant tax rates $\tau_t^w$ and $\tau_t^c$.

What is a default in this model? In period $t$, government has debt $DG_t$, revenue $\tau_t C_t + \tau_t w_t L_t$, non-financial expenses $CG_t + Z_t$ and needs to pay the interest on debt $DG_t$. Government will default on payments if the shock to expenditures $CG_t$ is so large that its revenue is too small to pay all expenses and it can no longer borrow; it will default if next period debt $DG_{t+1}$ would exceed the maximal level that investors tolerate, namely the debt capacity level $DG_{t+1}^s$, equal to the net present value of future surpluses. In other
words, a default will occur at period $t$ if and only if

$$DG_{t+1} > DG_{t+1}^s.$$  

Note that if the actual debt level $DG_t$ equals the debt capacity level $DG_t^s$ at time $t$ and there is no new information at time $t+1$, then by the government debt law of motion the actual debt level in the next period $DG_{t+1}$ equals the sustainable level $DG_{t+1}^s$; the laws of motion for $DG_t$ and $DG_t^s$ are expressed in a similar way. We thus have

$$DG_{t+1} > DG_{t+1}^s \iff (R_{t+1} + Pt)(DG_t - \mathcal{J}_t) > (R_{t+1} + Pt)(DG_t^s - E[\mathcal{J}_t])$$

$$\iff DG_t - \mathcal{J}_t > DG_t^s - E[\mathcal{J}_t] \iff E[\mathcal{J}_t] - \mathcal{J}_t > DG_t^s - DG_t.$$  

Since tax rates $\tau_c$ and $\tau_w$ are assumed to be constant, the interest rate schedule $r_t$ is given and consumption decisions $C_t$ are taken in expected values, aggregate labor supply decisions $L_t$, capital stock $K_t$, output $Y_t$ and wages $w_t$ do not fluctuate. By the law of iterated expectations then,

$$E[\mathcal{J}_t] - \mathcal{J}_t = CG_t + Z_t - \tau_c C_t - \tau_w w_t L_t - (E[CG_t] + Z_t - \tau_c C_t - \tau_w w_t L_t)$$

$$= CG_t - E[CG_t].$$

So there is default if and only if

$$CG_t - E[CG_t] > DG_t^s - DG_t \iff CG_t > DG_t^s - DG_t + E[CG_t].$$  

There is a default if the difference between the realized government expenditure (shock $CG_t$ observed at time $t$) and its average value ($E[CG_t]$) is larger than the difference between the debt capacity level and the actual debt level ($DG_t^s - DG_t$), this latest difference representing how much more debt can be borrowed by government. This is consistent with intuition. For instance, if government borrowed up to the maximum level, so $DG_t = DG_t^s$, any expenditure $CG_t$ which turns out to be larger than the expected value $E[CG_t]$ will lead to a default, since the debt capacity level is the maximum level of debt that government can borrow with the current tax revenue to cover future expected expenditures.

Recall that scaled shocks to government expenditures $CG/Y$ are distributed according to $\phi(E[CG_t]/Y, s^2)$, so the probability of default is equal to

$$\hat{p}_t = P(CG_t > DG_t^s - DG_t + E[CG_t])$$

$$= 1 - P\left(\frac{CG_t}{Y_t} \leq \frac{DG_t^s - DG_t + E[CG_t]}{Y_t}\right)$$

$$= 1 - \phi\left(\frac{DG_t^s - DG_t + E[CG_t]}{Y_t}\right).$$
Summarizing, we have obtained

**Lemma 4. (Probability of default):** assuming independent identically distributed shocks to government expenditures $CG/Y$ with cumulative distribution function $\phi(\overline{CG}/Y, \bar{s}^2)$, given interest rates $r_t$ and constant taxes and transfers $\tau_c$, $\tau_w$, $Z_t$, the probability of government defaults on debt repayment at time $t$ is given by

\[
\bar{p}_t = P(DG_t > DG^s_t) = P(CG_t - \overline{CG} > DG^s_t - DG_t) = 1 - \phi \left( \frac{DG^s_t - DG_t + \overline{CG}}{Y_t} \right)
\]

I now turn to the existence and a characterization of the public debt premium function, assuming the existence of a stationary steady-state. The following is a noteworthy consequence of the assumption, useful for the next and other results: the stationary steady-state level of actual debt equals the debt capacity level, $E[DG] = DG^s$. There is no room for Ponzi games. Indeed, we have by definition $E[DG] \leq E[DG^s] = DG^s$. Could the expected debt level be strictly greater, that is $E[DG] > DG^s$? If this was the case, the intertemporal budget constraint $DG_{t+1} = (R_{t+1} + P_t)(DG_t - S_t)$ for government debt would be violated, since the debt capacity level $DG^s$ is nothing but the net present value of future expected surpluses $E[S]$. The public debt premium function is characterized in the following way:

**Proposition 1. (Endogenous public debt premium):** assuming the existence of a stationary steady-state, constant taxes and transfers $\tau_c$, $\tau_w$, $Z_t$, independent identically distributed shocks to government expenditures $CG/Y$ with a cumulative distribution function $\phi(\overline{CG}/Y, \bar{s}^2)$ and perfect competition in the debt markets, then, given the inherited debt level $DG$, the interest rate premium on public debt in the steady-state is equal to

\[
\pi \left( \frac{DG^s}{Y} \right) = Y \frac{DG^s}{DG^s \bar{\delta}} = Y \frac{DG^s}{DG^s} \left[ 1 - \phi \left( \frac{DG^s - DG + E[CG_t]}{Y} \right) \right] \bar{s}
\]

**Proof:** because the proposition considers stationary steady-state, I drop time indices. First note that in the stationary steady-state, $E[DG] = DG^s$. Second, in perfectly competitive insurance markets, the costs of government defaults must equal the net premium revenue (in expectation). The costs of government defaults equal the probability $\bar{p}$ of default times the average size of a default. As shown by lemma 4, a default occurs if and only if $CG - E[CG_t] > DG^s - DG$. Under the assumptions used, we have that $E[DG^s] = DG^s = E[DG]$. The average size of a default equals the standard deviation of the stochastic process $CG$. Indeed, when
the realization $CG$ is smaller or equal to the average value $E[CG_t]$, the government is able to meet its commitments, since $E[DG^s] = E[DG]$ and the sustainable debt level $DG^s$ is the discounted future fiscal surpluses, taking average government expenditures $E[CG_t]$ as reference. When the realization $CG$ is larger than $E[CG_t]$, the difference between $CG$ and $E[CG_t]$ represents in average the excess commitments that the government cannot repay. The average difference between $CG$ and $E[CG_t]$ is nothing but the standard deviation of $CG$. Since $CG/Y$ has variance $\bar{s}^2$, the standard deviation of $CG$ equals that of $Y(CG/Y)$, namely $\sqrt{Y^2\bar{s}^2}$ or $Y\bar{s}$.

The expected net premium revenue equals the gross revenue, that is the premium rate $\pi(DG^s/Y)$ times the average loan size. In a stationary steady-state, the average loan size is the government debt level $E[DG]$, which, by definition of the debt capacity level as an expectation, equals $E[DG^s] = DG^s$. Because capital markets are integrated, investors can lend to government in different countries. As expenditure shocks are uncorrelated, investors lend to different governments to diversify risk. One can thus apply a law of large numbers.

Equating net expected costs and revenues and using the law of large numbers, we have

$$\bar{p}Y\bar{s} = \pi\left(\frac{DG^s}{Y}\right)DG^s$$

or, using lemma 4, the result:

$$\pi\left(\frac{DG^s}{Y}\right) = \frac{Y}{DG^s} \bar{p} \bar{s} = \frac{Y}{DG^s} \left[ 1 - \phi\left(\frac{DG^s - DG + E[CG_t]}{Y}\right) \right] \bar{s}$$

QED.

The proof of proposition 1 shows that the public debt premium $\pi(DG^s/Y)$ is nothing but an insurance premium, under the assumptions of perfect competition in the debt markets and independence and identical distribution of shocks to government expenditures. These two assumptions are key so worth a discussion.

Debt premiums being insurance premiums has other consequences. First, premiums are collected into a separate fund and are priced at the lowest possible level such that the fund covers defaults, without any profit. The premium is no more no less than the average cost of defaults. As a consequence, there is no arbitrage nor portfolio allocation question and the return to households savings remains the deterministic $R_t = 1 + r_t$. The insurance isolates households from the stochastic nature of defaults. Second, because shocks are independent and capital markets integrated, insurance premiums have a cross-sectional dimension: investors lend to different governments to diversify risk.

Third, perfect competition guarantees that there are no multiple equilibrium. Assuming the existence of a well-behaved premium function $P = \pi(DG^s/Y)$, in the sense that the premium declines with capacity ($\pi' < 0$), there is no premium for infinite capacity ($\pi(0) = \infty$) and the premium is infinitely high for no capacity ($\pi(\infty) = 0$), is not sufficient
to rule out multiple equilibriums. Indeed, the sole definition (3) of debt capacity
\[
DG^t_i = \sum_{u=t}^{\infty} \left( \prod_{v=t}^{u-1} \frac{1}{R_{v+1} + P_v} \right) E[U_u] = \sum_{u=t}^{\infty} \left( \prod_{v=t}^{u-1} \frac{1}{R_{v+1} + \pi(DG^t_v/Y_v)} \right) E[U_u] \tag{6}
\]
and existence of a well-behaved premium \( P = \pi(DG^s/Y) \) lead to several possible equilibriums. One equilibrium is the point \( DG^s = 0 \), associated with \( \pi(0) = \infty \). Another equilibrium is the point \( DG^s \) associated with the premium \( \pi(DG^s/Y) \) given by proposition 1. The definition of the debt capacity and the existence of a well-behaved premium function can thus lead to multiple equilibrium. Adding the condition of perfect competition in the debt markets however rule them out. Indeed, given predictions of expected fiscal surpluses \( E[S_u] \), the government will simply choose the offer from the investor with the lowest premium \( P \). Note that the offers are strictly bounded from below by zero \( (P > 0) \), so the notion of the lowest premium is well defined: the premium is indeed nothing but an insurance premium and the risk of default is strictly greater than zero.

Are the two key assumptions on competition and shocks plausible? In reality, none of the two hold perfectly but the higher the number of countries considered in the same currency area, the more likely they do\(^{18} \).

Next I derive some characteristics of the endogenous premium function \( \pi \). I denote by \((H2)\) the condition on the shock distribution \( \phi \) that \( \phi'' \) is either negative or small enough that \(- \frac{Y}{DG^s} \left( \frac{1}{DG^s} (1 - \phi'(x)) + \phi''(x) \right) < 0 \ \forall x \). There are several distributions \( \phi \) which satisfy condition \((H2)\). For instance, if \( \phi \) is uniform, then \( \phi'' = 0 \); another example are exponential distributions, since \( \phi'(x) = \lambda e^{-\lambda x} \) and \( \phi''(x) = -\lambda^2 e^{-\lambda x} < 0 \).

**Proposition 2. (Properties of the premium function):** assuming the existence of a stationary steady-state, constant taxes and transfers \( \tau_c, \tau_w, Z_t \), independent identically distributed shocks to government expenditures \( CG/Y \) with cumulative distribution function \( \phi(CG/Y, \bar{s}^2) \) and perfect competition in the debt markets, then the endogenous interest rate premium function on public debt in the long run is declining in the debt sustainable level and, in some cases, convex:

\[
\frac{\partial}{\partial DG^s_Y} \pi \left( \frac{DG^s}{Y} \right) < 0 \quad \text{if condition (H2) holds:} \quad \frac{\partial^2}{\partial (DG^s_Y)^2} \pi \left( \frac{DG^s}{Y} \right) > 0
\]

\(^{18}\text{Asymmetric information and transaction costs make the debt market imperfect, allowing investors to make profits. Assuming a perfect market however provides results which are useful benchmarks, in the same way that macroeconomic theories assume that firms make no profit. The international component of business cycles introduces a correlation into shocks to government expenditures. Shocks may also have different distributions due to differences in social norms or climate exposure. However, distribution differences between countries may not be very big in common currency areas. For large enough currency areas and considering different debt maturities, uncorrelated shocks and risk diversification opportunities are more realistic.}\)
The proof follows from differentiation of the premium function given by proposition 1. Calculations can be found in the technical appendix Davoine (2013).

The fact that \( \pi\left(\frac{DG^s}{Y}\right) \) is a declining function means that the risk premium decrease with the level of debt sustainability, an intuitive feature. A low level of debt capacity \( DG^s \), relative to the size \( Y \) of the economy, tells an investor that a government is not in a position to generate a lot of surplus, a signal of higher probability of default. Investors will thus charge a higher premium \( P = \pi\left(\frac{DG^s}{Y}\right) \) if \( DG^s \) is low. This is also consistent with the intuition one can form from household loans: a bank is prepared to make a loan to a household up to the point that the household can repay the loan, taking into account future income and expenses streams; for the same loan amount and expenditures, the bank will charge a lower premium to a household with large revenues than to a household with low revenues.

The fact that, in some cases, the function is convex means that the lower the level of debt capacity, the higher the increase in the risk premium. This, again, is intuitive: a government running a low revenue tax schedule and exposed to high expenditure shocks will have by definition a low level of debt capacity; when the capacity level is very low, the risk of default is very high, ceteris paribus, so investors will charge a very large premium. Ardagna, Caselli, and Lane (2007) provide empirical evidence supporting the intuition and the result, finding that government bond rates increase more than proportionally with the debt-to-GDP ratio.

Quantitative analysis in section 4 will illustrate the importance of convexity.

Proposition 2 only applies to debt capacity \( DG^s \) but can be adjusted to actual debt \( DG \). Differentiation of (5) shows that the premium increases with the inherited debt stock \( DG \), consistent with intuition and the textbook result from Eaton and Gersovitz (1981). It increases in a convex manner if and only if \( \phi'' \leq 0 \), which holds in particular for uniform and exponential \( \phi \) shock distributions.

I now perform simple investigations of the impact of policy reforms on public debt capacity. The goal is not to obtain general results but to build intuition and develop an understanding of the mechanisms at work. Some of the results simply confirm the intuition that one can develop from inspection.

One would expect from inspection of (6) that a decrease in the premium \( P_t \) or increase in the average fiscal surplus \( E[\mathcal{S}] \), whatever their reasons, would lead to an increase in the debt capacity. The cheaper the cost of public debt or the higher the income stream, the easier it is for government to repay high debt and the higher the debt they can afford in a sustainable fashion.

The following result confirms these intuitions, and more, in a simple setting.

**Lemma 5.** *(Policy impacts on public debt)* assume conditions of proposition 1, \( DG^s > 0 \) and a constant interest rate \( R \). Then policy impacts debt capacity \( DG^s \) in the following way:
if consumption taxes $\tau_c$ are small, exogenous changes in the premium $P = \pi(DG^s/Y)$ lead to
\[
\frac{\partial DG^s}{\partial P} < 0.
\]

if the premium derivative $\pi'(DG^s/Y)$ is small (in absolute value),
\[
\frac{\partial DG^s}{\partial E[S]} > 0.
\]

if the current policy is on the government revenue increasing side of the Laffer curve ($\tau_w L$ increases if $\tau_w$ increases) and output variations $\partial Y$ are small,
\[
\frac{\partial DG^s}{\partial \tau_w} > \frac{\partial DG^s}{\partial \tau_w} \bigg|_{\mathscr{S} = \text{const}} > 0
\]

and similarly for other fiscal consolidation efforts, $\frac{\partial DG^s}{\partial \tau_c}$ and $\frac{\partial DG^s}{\partial E[CG]}$.

Note that the condition that the premium derivative $\pi'(DG^s/Y)$ is small holds if the premium $\pi$ function is convex decreasing (as for instance under conditions identified in proposition 2) and the initial debt capacity $DG^s/Y$ is large. The proof of lemma 5 is mechanical differentiation and re-arrangement of terms.

**Proof:** in a stationary steady-state and with constant interest rate, the expression of public debt capacity can be simplified:
\[
DG^s = \sum_{u=t}^{\infty} \left( \prod_{v=t}^{u-1} \frac{1}{R_{v+1} + P_v} \right) E[S_u] = \left\{ \sum_{u=t}^{\infty} \left( \frac{1}{R + P} \right)^{u-1} \right\} E[S_u] = \frac{R + P}{R + P - 1} E[S]
\]

With constant interest rate and fiscal policy, households labor supply is unchanged. When consumption taxes are small, $\frac{\partial E[S_u]}{\partial \tau_c} = \frac{\partial}{\partial \tau_c} (\tau_c C_u + \tau_w w_u L_u - (E[CG] + Z_u)) \equiv 0$. Thus
\[
\frac{\partial DG^s}{\partial P} = \frac{\partial}{\partial P} \frac{R + P}{R + P - 1} E[S] \cong \left\{ \frac{\partial}{\partial P} \frac{R + P}{R + P - 1} \right\} E[S] = -\left( \frac{1}{R + P - 1} \right)^2 E[S] < 0.
\]

Next,
\[
\frac{\partial DG^s}{\partial E[\mathscr{S}]} = \frac{\partial}{\partial E[\mathscr{S}]} \frac{R + P}{R + P - 1} E[S] = \left\{ \frac{\partial}{\partial E[\mathscr{S}]} \frac{R + P}{R + P - 1} \frac{\partial}{\partial E[\mathscr{S}]} \right\} E[S] + \frac{R + P}{R + P - 1}
\]
\[
= - \left\{ \frac{\partial}{\partial E[\mathscr{S}]} \right\} \left( \frac{1}{R + P - 1} \right)^2 E[S] + \frac{R + P}{R + P - 1}.
\]

We have, from proposition 1, that the premium can be expressed as a function of debt
capacity, so \( \frac{\partial P}{\partial E[\mathcal{S}]} = \frac{\partial}{\partial E[\mathcal{S}]} \pi \left( \frac{DG_s}{Y} \right) = \pi' \left( \frac{DG_s}{Y} \right) \frac{1}{Y} \frac{\partial DG_s}{\partial E[\mathcal{S}]} \). Combining,

\[
\frac{\partial DG_s}{\partial E[\mathcal{S}]} = -\pi' \left( \frac{DG_s}{Y} \right) \frac{1}{Y} \frac{\partial DG_s}{\partial E[\mathcal{S}]} \left( \frac{1}{R + P - 1} \right)^2 E[\mathcal{S}] + \frac{R + P}{R + P - 1} \frac{R + P}{1 + \pi' \left( \frac{DG_s}{Y} \right) \frac{1}{Y} \left( \frac{1}{R + P - 1} \right)^2 E[\mathcal{S}] R + P - 1.}
\]

When \( \pi'(DG_s/Y) \) is small in absolute value, \( \frac{\partial DG_s}{\partial E[\mathcal{S}]} > 0 \). From this expression and the proposition 2 result that \( \pi'(DG_s/Y) < 0 \) follows

\[
\frac{\partial DG_s}{\partial E[\mathcal{S}]} > \frac{\partial DG_s}{\partial E[\mathcal{S}]} |_{P=\text{const}} = \frac{R + P}{R + P - 1} > 0.
\]

Noting that \( \frac{\partial DG_s}{\partial \tau_w} = \frac{\partial E[\mathcal{S}]}{\partial \tau_w} \left\{ \frac{\partial DG_s}{\partial E[\mathcal{S}]} \right\} \) and that \( \frac{\partial E[\mathcal{S}]}{\partial \tau_w} > 0 \) on the revenue increasing side of the Laffer curve, concludes.

QED.

The first part of lemma 5, \( \partial DG_s/\partial P < 0 \), confirms the intuition that the public debt capacity increases when the premium decreases. The second part, \( \partial DG_s/\partial E[\mathcal{S}] > 0 \), confirms that the debt capacity increases when the average fiscal surplus increases. The last part of the lemma identifies a multiplier effect. The increase in debt capacity coming from a fiscal consolidation effort leads not only to an increase in the fiscal surplus, which increases the debt capacity alone, but also to a decline in the premium which further increases the debt capacity: the total increase \( \partial DG_s/\partial \tau_w \) is larger than the increase due to taxes alone, holding the premium constant, as \( \frac{\partial DG_s}{\partial \tau_w} |_{P=\text{const}} \). The discussion gives a hint of possible advantages of fiscal consolidation efforts over the introduction of Eurobonds, for countries which are struggling with low sustainability of their public finances and paying high premiums on their debt. Eurobonds have a direct effect, as they reduce the premium and increase the sustainable level of public debt. Fiscal consolidation on the other hand can benefit from the multiplier effect, as they both directly improve the average fiscal surplus and then reduce the premium. The key is the endogeneity of the public debt premium, defined in terms of the debt capacity. The higher the debt capacity, the lower the risk of a sovereign default and the lower the insurance premium. Quantitative analysis in section 4 will provide an illustration of the multiplier benefit of fiscal consolidations in a general equilibrium context.

### 3.3 Intertemporal trade benefits

In this section I analyze the welfare benefits of intertemporal trade, compared to autarky. The aim of the section is to generalize some of the results that Buiter (1981) obtained for 2 countries with no government, exogenous labor supply and where the only difference
are time preferences. Buiter (1981) shows that moving to trade is not always beneficial, that the interest rate in trade is strictly comprised between the autarky interest rates of the two countries and that the country with more impatient households will be net importer of goods (international borrower).

I obtain similar results, where in some cases opening to trade is beneficial and in other cases, it is not. I also provide a quantitative criteria to differentiate the cases.

Denote by \((Hi)\) the following assumption for country \(i\): utility is logarithmic \((\sigma^i = 1)\), mortality is low \((\gamma^i \text{ close to } 1)\), consumption taxes are low \((\tau^i_c \text{ small})\); there are no Eurobonds \((\mathcal{E}^E = 0)\); welfare is positive \(Q^i > 0\); government has low debt \(DG^i > 0\) and \(\pi(.)\) is convex decreasing (enough that \(\partial DG / \partial r\) is small).

**Lemma 6. (Welfare gains from intertemporal trade):** let \(M\) countries have autarky equilibrium interest rates \(r^1_{NT}, r^2_{NT}, \ldots, r^M_{NT}\). Then in the long run opening to trade will lead to a common interest rate \(r\). Each country \(i\) such that \(r^i_{NT} > r\) will initially run a negative trade balance (trade deficit), accumulate a negative foreign asset position and have to switch to a positive trade balance (trade surplus) in the long run to pay for the foreign debt, and vice-versa for each country \(i\) such that \(r^i_{NT} < r\). Finally, if conditions \((Hi)\) are satisfied,

\[
\text{Country } i \text{ will benefit from trade in the long run} \iff T_1^i T_2^i (r - r^i_{NT}) > 0
\]

where

\[
T_1 = \frac{1 + r}{(1 + \tau_c)\Omega - \frac{1 + r}{r}(1 + \pi_h)}
\]

\[
T_2 = -\frac{1 + \pi_h}{1 + r} C - (CG + Z - \tau_w w L) + NFL \left( \frac{L}{N} \right) - \pi_h (Y - \delta K)
\]

The proof is lengthy and tedious but not conceptually informative. It can be found in Davoine (2013). The same document contains a discussion of this result.

### 4 Quantitative results

In this section I numerically compare the effect of Eurobonds to well-known policy reforms, namely tax reforms and cross-country transfers. I focus on steady-state outcomes and briefly mention transition effects.

As noted in the introduction, there are only few numerical studies of intertemporal trade effects in open economy settings, allowing to isolate the effect of differences in public
4. QUANTITATIVE RESULTS

policy. My approach is closest to that of Jäger and Keuschnigg (1991). Essentially, my analysis is an extension to endogenous labor supply, endogenous public debt premium, cross-country transfers and Eurobonds.

4.1 Model and calibration

For smooth transitions and more realistic predictions, I add capital adjustment costs, lump sum taxes and trade costs to the model (see appendix A). The discussion will only make references to the basic model. With one exception, the calibration and parameters are standard. Only the specification and parameters for the public debt premium are not frequent (see appendix B). Numerical tools, based on the Fair and Taylor (1983) algorithm, are described in the technical appendix Davoine (2013).

I compare outcomes for two countries, 1 and 2, which are identical except that they have different tax and spending policies. In particular, households have identical preferences and production uses the same technology. Specifically, country 1 relies more on consumption taxes and has a larger average government consumption. In the base case, parameter choices for the public debt premium function and policy variables are arbitrarily made such that the public debt capacity is low in country 1 while it is close to the Maastricht solvency criteria for country 2, the premium in each country being consistent with the observed spectrum displayed in figure 1:

<table>
<thead>
<tr>
<th>Policy choice</th>
<th>Country 1</th>
<th>Country 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor income tax rate $\tau_w$</td>
<td>0.15</td>
<td>0.20</td>
</tr>
<tr>
<td>Consumption tax rate $\tau_c$</td>
<td>0.05</td>
<td>0.01</td>
</tr>
<tr>
<td>Government Consumption $E[CG]$</td>
<td>12</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Country 1</th>
<th>Country 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resulting Public Debt premium $\pi(DG^w/Y)$</td>
<td>18%</td>
<td>0.03%</td>
</tr>
<tr>
<td>Resulting Public Debt Capacity $DG^w/Y$</td>
<td>7%</td>
<td>69%</td>
</tr>
</tbody>
</table>

Autarky equilibrium allocations in this base case differ by country. Country 1 has a lower interest rate. Indeed, the household saving decisions are defined by the optimality condition $Q^i = (V^i + DG^i + DF^i + H^i)/(1 + \tau^i_i(1 + \Omega^i_i))$. Ceteris paribus and compared to country 2, higher consumption taxes in country 1 push consumption down, savings and capital accumulation up and thus the interest rate down. Interest rates in autarky are 5.4% in country 1 and 6.0% in country 2. At trade opening, households in country 1 are

---

19One can think of countries 1 and 2 as two different countries or as two different sets of countries of the same currency area. One can also focus on country 1 and consider that country 2 represents the rest of the currency area. The only difference is the relative size of the countries.
attracted by the high interest rate in country 2, send part of their savings in that country until the interest rates equalize, at 5.8%. Country 2 ends up with a negative foreign asset position.

4.2 Numerical experiments

I investigate the effect of changes in fiscal policies in country 1, cross-country transfers or implementation of Eurobonds. I will consider effects on standard macroeconomic outcomes, including output, consumption and labor supply, as well as welfare and public finances. When considering long run outcomes, \( Q = C - N \varphi (L/N) \) is an appropriate measure of welfare\(^{20}\).

I assume the existence of stationary steady-states and that governments are committed to constant public debt in steady-states and to vary fiscal policy during transitions to avoid explosive debt paths.

Given fiscal policy, the expected level of the debt is defined endogenously in any steady-state equilibrium. In general, one degree of freedom needs to be removed to obtain the equilibrium. Quantitative analysis often choose a debt level and let one fiscal instrument adjust endogenously\(^{21}\). I do the opposite: I fix the value of all fiscal instruments and let the debt level adjust endogenously. I use this approach because the impact of Eurobonds on public debt is interesting and thus an endogenous level of debt important.

In details, the endogenous debt level is defined in the following way. As noted in section 3.2, expected public debt equals the debt capacity, \( E[DG_i] = DG^s \) when one assumes the existence of a stationary steady-state. The exogenous fiscal policy uniquely defines the expected fiscal surplus \( E[\mathcal{S}] = \tau_c C + \tau_w w L - (E[CG] + Z) \): capital stock \( K \) and labor supply decisions \( L \) are directly defined by the interest rate and fiscal policy; output \( Y \) and investment \( I \) follow; the interest rate, taken as given by each country in the multi-country setting, then defines excess demand of domestic goods (trade balance \( Y - C - I - E[CG] - Z \)), fixing aggregate consumption \( C \) and thus the expected surplus \( E[\mathcal{S}] \).

In a stationary steady-state, both the interest rate and the premium on the public debt are constant, leading to a simplified expression for public debt capacity,

\[
DG^s = \sum_{u=1}^{\infty} \left( \prod_{v=1}^{u-1} \frac{1}{R + \pi (DG^s/Y)} \right) \frac{R + \pi (DG^s/Y)}{R + \pi (DG^s/Y) - 1} E[\mathcal{S}].
\]

Altogether the public debt capacity level is thus (implicitly) uniquely defined\(^{22}\). This equation illustrates the difference between my approach and other quantitative studies:

\(^{20}\)See footnote 17.

\(^{21}\)Typically the post-reform debt level is assumed to be identical to the pre-reform level. One refers to the chosen fiscal instrument, such as labor income taxes or government consumption, as the closing policy instrument for the government budget constraint.

\(^{22}\)Lemma 7 in appendix C presents details and expressions after substitution of steady-state values.
to remove one degree of freedom, I fix fiscal policy and thus $E[S]$, which then defines $DG^s$. In contrast, other quantitative studies fix the debt level $DG^s$ and let one fiscal policy instrument adjust (for instance $\tau_w$), defining $E[S]$.

The comparative static experiments consist in fixing one fiscal policy in each country before and after the reforms, computing the corresponding initial and final steady-state equilibrium, which includes the (endogenous) levels of public debt capacity. For instance, the labor income tax in country 1 is increased 10%. Macroeconomic outcomes as well as public debt capacity before and after the reform are then compared. The same approach is used to evaluate Eurobonds, leaving initial fiscal policy unchanged and comparing equilibrium with no Eurobonds, as initial steady-state, to equilibrium with Eurobonds, as final steady-state. In this approach, there is no need to choose a closing policy instrument for the government budget constraint.

Experiments with transition path outcomes consist in the following extension. Initial and final steady-state equilibrium are derived in the same way as comparative static experiments. The initial and final levels of public debt are endogenously defined. The path of debt between these two levels is however not unique and may not converge to the final debt level. To rule out explosive debt paths, I assume that government adjusts one tax rate or its average consumption so that the public debt level follows an arbitrary (quadratic-shaped) path between the initial and final steady-stage levels\(^{23}\). In contrast to the comparative statics experiments, a choice needs to be made for closing the government budget constraint, between the various tax rates and government consumption. The choice may impact the outcome, as some instruments are distortionary and some are not. I will discuss this choice in transition path experiments.

Most of the policy reform experiments that I conduct are equivalent to one-time tax cuts financed by later improvements in fiscal surpluses. For instance, the experiment of a long-run increase in income taxes $\Delta \tau_w > 0$ in country 1 is equivalent to one-time labor income tax cuts, leading to increased public debt financed by subsequent higher labor income taxes, investigated for instance by Persson (1985), Frenkel and Razin (1986) as well as Jäger and Keuschnigg (1991). In contrast to these studies, I assume that the price of public debt varies.

### 4.3 Expectations

What effects can we expect from tax reforms, cross-country transfers and Eurobonds? In this section, I collect answers from the theoretical and quantitative literature, as well as analytical results from section 3. These answers will be useful to analyze the results from the policy reform simulations, presented in continuation.

Due to general equilibrium, even simple policy reforms lead to complex effects. Complexity is increased in multiple countries environments and can lead to surprising results.

\(^{23}\text{This procedure is similar to that of Jäger and Keuschnigg (1991).}\)
Persson (1985) adds government debt to the Buiter (1981) model and analytically compares a one-time cut in taxes in single and multiple country cases. The tax cut increases debt, financed by a later tax increase, so the reform redistributes towards old generations. In single economies, this redistribution takes place at the expense of future generations. In some cases under open trade however, future generations can also benefit from the reform, in spite of higher taxes. If the country is a strong net lender on the international capital market, the increase in debt crowds out capital markets and increase the interest rate, so that future generations in foreign countries pay a very high price for their debt, to the benefit of future domestic generations. International redistribution dominates intergenerational redistribution for future home generations. Cross-country transfers may also lead to paradoxes. As shown by Galor and Polemarchakis (1987) and Haaparanta (1989), transfers may increase welfare of the donor country and decrease welfare of the recipient country, for reasons which I will discuss at a later stage.

Quantitative analysis also leads to surprising outcomes, even in environments with realistic calibrations. Mendoza and Tesar (1998) assess the welfare and spillover effects of replacing capital income taxes in the United States by either labor income or consumption taxes, assuming open trade with Europe and keeping public debt constant. Baseline results are consistent with the supply-side economics analysis of Lucas (1990), welfare gain in the US and welfare loss in Europe. However, these results are not robust: when foreign income is taxed, a realistic possibility, results are overturned. Using the Buiter (1981) model with a richer overlapping-generations structure, Jäger and Keuschnigg (1991) analyze the effect of public debt variation and generally find expected results, but not always. In particular, they “report some surprising results: first, for plausible parameter constellations, increases in public debt can increase the welfare of future generations. Second, allowing for asymmetric parameter constellations, all kinds of long-run welfare effects seem to be possible: utility of future generations may either fall or increase in both countries, or they may be affected in opposite ways.” (p.198).

A number of channels have been identified by the literature to explain these domestic and spillover effects of public policy reforms, most of which apply here. To discuss these channels and how they relate to the model, it is useful to recall optimality conditions for household decisions (see lemma 3 and its discussion),

\[ Q = \frac{1}{(1 + \tau_c)\Omega} (V + DG + DF + H) \]

Changes in household wealth \( V + DG + DF + H \) characterize changes in welfare \( Q = C - \varphi(L) \) and consumption \( C \), a convenient way to capture intertemporal effects on households consumption decisions. For instance, decreases in labor income taxes \( \tau_w \) push labor supply \( l \) up as well as human wealth \( H \), the net present value of future labor income streams, ceteris paribus. As a result, households have higher total wealth and can afford more consumption \( C = Q + \varphi(L) \).

I now present five basic domestic and spillover effects of public policy reforms: crowding-
out, related to the firm value component $V$ of household wealth; twin-deficits, related to foreign assets $DF$; intertemporal terms-of-trade effects, also related to $DF$; human wealth effects, related to $H$; and the multiplier effect, related to public debt $DG$. For simplicity of the exposition, I assume that the public policy reform leads to an increase in public debt, such as an increase in government expenditures.

**Crowding-out effect**  As discussed by Persson (1985), an increase in public debt creates an excess demand in the capital market, pushing up the equilibrium interest rate.

**Twin-deficits effect**  In a small open economy, the increase in public debt also creates an excess demand in the capital market, which is met by an increase in borrowing from foreign countries. The interest rate is indeed constant and domestic capital formation is thus also constant. The fiscal deficit increase is associated with a trade deficit increase. In contrast, excess demand is entirely met by domestic capital and a large increase in interest rate in a closed economy (crowding-out). As noted by Persson (1985), the multi-country case is an intermediate between the small open and closed economy cases, so both the twin-deficit and crowding-out effects take place.

**Intertemporal terms-of-trade effects** 24 Because of the twin-deficits effect, an increase in public debt leads to a decrease in foreign assets. On the other hand, the crowding-out effect pushes interest rates up. Both effects work on opposite directions on the absolute interest payments on foreign debt. As shown by Frenkel and Razin (1986), the net welfare effect over the long run is ambiguous and depends, in particular, on the initial position on the international capital market. For instance, if the home country is a net international lender, decides to cut taxes, remains a net lender after the reform and the crowding-out effect is strong, then the interest rate rises much, the home country benefits from high debt service payments by foreign countries, which can dominate all other welfare effects. In such a case, the cut in taxes and rise in public debt not only benefits domestic current generations, but also future domestic generations, thanks to redistribution between countries (Persson, 1985).

**Human wealth effect**  Crowding-out leads to increased interest rate with higher public debt. In a general equilibrium environment and as noted by Haaparanta (1989) for instance, this leads to lower equilibrium wages. Human wealth, the net-of-tax present value of future labor income, is also pushed down (ceteris paribus, in particular at constant labor supply).

---

Multiplier effect  This effect is new and derived from propositions 1 and 2 as well as lemma 5. Reforms which modify fiscal surpluses have a direct impact on public debt capacity. This modifies the interest rate premium, magnifying the initial impact on public debt: an increase in $E[\mathcal{G}]$ has a direct impact on capacity $DG_t^s = \sum_{u \geq t} \frac{1}{Ru + \pi(DG_y^s/Y_u)} E[\mathcal{G}_u]$ and an indirect impact, since the premium $\pi(DG^s/Y)$ declines with capacity $DG^s$. Reforms which modify the public debt capacity via the premium do not benefit from this multiplier effect (see section 3.2 for a longer discussion).

Most, if not all, of these effects play a role in policy reforms. In general, they do not work in the same direction and it is not easy, if not impossible, to anticipate their net welfare effect in each country. The convexity of the public debt premium function will play an important role in explaining net effects. Numerical simulations on the other hand provide clear answers, even if they do not say why. Giving an intuitive account of each separate effect is feasible but providing an intuitive account for the net effect remains a difficult task, as can be acknowledged by the fact that intuitive accounts for the net effect is most of the time ignored in existing numerical studies of intertemporal trade effects.

4.4 One-dimensional policy reforms

Given current fiscal policy, is it preferable to change taxes or introduce a Eurobond? In this section I compare reforms involving a single policy dimension: introduction of a Eurobond, fiscal consolidation, small and large cross-country transfers. I present and discuss labor income tax increases as consolidation measure. Consolidation with consumption tax increases or government expenditures decreases lead to similar results. I present steady-state results and focus on the main outcomes. The next section considers transition results.

Table 1 provides the results. The introduction of a Eurobond (column $E = 1$) is compared to a fiscal consolidation measure (column $\Delta \tau_w > 0$) and small transfers (column $Z_1 = -0.5$) such that they each achieve the same public debt capacity level$^{25}$. It is also compared to a large transfer (column $Z_1 = -1$).

There are several noteworthy outcomes. First, all policies raise the debt capacity level of country 1, as expected. Some preserve the debt capacity of country 2, but not all. Second, fiscal consolidation measures have a positive impact on welfare in country 1. Third, small transfers also have positive impact on welfare and consumption in the recipient country (1). Fourth, large transfers have a negative impact on welfare in the recipient country (1) and a positive impact in the giving country (2). Fifth, Eurobonds have a negative impact on welfare in country 1 and a positive impact in country 2.

$^{25}$To give a sense of the shocks, tax reforms amount to an increase of about 0.6% of GDP in government tax revenue. Absolute variations are provided at the bottom of table 1 for the tax reform instrument. Other fiscal policy instruments are not changed in the long-run steady-state, as budget closing instruments are only useful for transition path outcomes (see section 4.2).
4. QUANTITATIVE RESULTS

<table>
<thead>
<tr>
<th>Country</th>
<th>Base</th>
<th>$\mathcal{E} = 1$</th>
<th>$\Delta \tau_w &gt; 0$</th>
<th>$Z_1 = -0.5$</th>
<th>$Z_1 = -1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest rate $r$</td>
<td>5.79%</td>
<td>5.73%</td>
<td>5.80%</td>
<td>5.76%</td>
<td>5.82%</td>
</tr>
<tr>
<td>Wage rate $w$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labor supply $L$</td>
<td>1</td>
<td>3.0</td>
<td>0.2%</td>
<td>0.0%</td>
<td>0.1%</td>
</tr>
<tr>
<td>Capital $K$</td>
<td>1</td>
<td>215</td>
<td>0.6%</td>
<td>-0.3%</td>
<td>0.3%</td>
</tr>
<tr>
<td>Output $Y$</td>
<td>1</td>
<td>100.0</td>
<td>0.2%</td>
<td>-0.3%</td>
<td>0.1%</td>
</tr>
<tr>
<td>Disp income $(1 - \tau_w)wL + \tau_{lump}$</td>
<td>1</td>
<td>56.1</td>
<td>+0.2%</td>
<td>-1.3%</td>
<td>+0.1%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>52.6</td>
<td>+0.2%</td>
<td>-0.0%</td>
<td>+0.1%</td>
</tr>
<tr>
<td>Trade Bal $(TB - TB_0)/Y_0$</td>
<td>1</td>
<td>-</td>
<td>+1.8%</td>
<td>-0.2%</td>
<td>+0.1%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-</td>
<td>-2.2%</td>
<td>-0.1%</td>
<td>-0.4%</td>
</tr>
<tr>
<td>Public Premium $\pi(DG^s/Y)$</td>
<td>1</td>
<td>17.7%</td>
<td>1.4%</td>
<td>5.2%</td>
<td>5.2%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.0%</td>
<td>1.4%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Pub Debt Capacity $DG^s/Y$</td>
<td>1</td>
<td>0.07</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.69</td>
<td>0.60</td>
<td>0.69</td>
<td>0.59</td>
</tr>
<tr>
<td>Firm value $V/Y$</td>
<td>1</td>
<td>2.28</td>
<td>2.29</td>
<td>2.28</td>
<td>2.28</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.28</td>
<td>2.29</td>
<td>2.28</td>
<td>2.28</td>
</tr>
<tr>
<td>Foreign Assets $DF/Y$</td>
<td>1</td>
<td>0.52</td>
<td>0.11</td>
<td>0.50</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-0.52</td>
<td>-0.11</td>
<td>-0.50</td>
<td>-0.45</td>
</tr>
<tr>
<td>Assets $(DG^s + V + DF)/Y$</td>
<td>1</td>
<td>2.87</td>
<td>2.59</td>
<td>2.97</td>
<td>2.93</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.45</td>
<td>2.75</td>
<td>2.47</td>
<td>2.42</td>
</tr>
<tr>
<td>Human wealth $H/Y$</td>
<td>1</td>
<td>5.97</td>
<td>6.00</td>
<td>5.90</td>
<td>5.98</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5.61</td>
<td>5.65</td>
<td>5.61</td>
<td>5.63</td>
</tr>
<tr>
<td>Consumption $C$</td>
<td>1</td>
<td>68.8</td>
<td>-2.5%</td>
<td>-0.0%</td>
<td>+0.7%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>65.3</td>
<td>+3.6%</td>
<td>+0.1%</td>
<td>-0.2%</td>
</tr>
<tr>
<td>Welfare $(Q - Q_0)/Y_0$</td>
<td>1</td>
<td>-</td>
<td>-1.7%</td>
<td>+0.1%</td>
<td>+0.5%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-</td>
<td>+2.3%</td>
<td>+0.1%</td>
<td>-0.1%</td>
</tr>
</tbody>
</table>

All policies are designed to reach the same debt capacity level, except the last column. Signed percentage numbers represent variations from the base case. Welfare and trade balance changes are normalized to remove scale effects. A “+” sign in the trade balance change means that the trade balance is more positive or less negative, and vice versa. $\mathcal{E} = 1$: introduction of a Eurobond; $\Delta \tau_w > 0$: labor income tax $\tau_w$ changes from 0.150 to 0.159 in country 1, equivalent to a 0.6% of GDP increase in tax revenue; $Z_1 = -0.5$: small transfer from country 2 to country 1; $Z_1 = -1$: large transfer from country 2 to country 1 equivalent to 1% of GDP.

Table 1: Eurobonds compared to other policy changes, over the long run

The first outcome is expected. Eurobonds raise the debt capacity level $DG^s$ of country 1 from 7% of GDP to 19% of GDP, as the premium is reduced from 17.7% to 1.4%. Fiscal consolidation measures and cross-country transfers increase the expected fiscal surplus $\mathcal{S} = \tau_c C + \tau_w wL - CG - Z$ and thus the debt capacity level. Because fiscal consolidations do not impact country 2 directly and general equilibrium effects are moderate (the interest rate increases from 5.79% to 5.80%), the debt capacity level of country 2 is unaffected by these reforms. Eurobonds on the other hand increase its public debt premium and transfers reduce its expected fiscal surplus. Both reduce the debt capacity of country 2, from 69% of GDP to 60% of GDP or lower.
In light of previous literature results (see section 4.3), the second outcome may or may
not be expected. In spite of negative impacts on output, fiscal consolidation increases
welfare in country 1. Increases in labor income taxes raise welfare\(^{26}\) by about 0.1%.
The characterization of optimal behavior given by (7) is useful to assess the welfare
impact of various policy reforms. I introduce this analytical tool in details for the labor
income tax reform and will later make use of the tool in a more direct way for other
reforms.

The increase in labor income tax is a small domestic shock and leads to small cross-
country effects and small change in factor prices. The interest rate increases slightly
at 5.80%. As expected, labor supply declines with the tax increase. Given the small
elasticity of labor supply however, the decline is moderate, at \(-0.2\%\). Crowding-out
increases demand for savings, so the interest rate increases, dragging down investment
and capital \(-0.3\%\). The supply of production factors declining, total output also
decreases \(-0.3\%\). While output changes provide one measure of the impact of the reform,
welfare impacts go beyond. The expression of welfare in terms of wealth components
\(V, H, DF\) and \(DG^s\) given in (7) is useful to assess welfare impacts. Drops in firm value
\(V\) and human wealth \(H\) follow drops in capital \(K\), labor supply \(L\) and disposable income
\((-0.5\) and \(-7\) percentage point of GDP, respectively). The decrease in fiscal deficit be-
ing associated with an immediate decrease in trade deficit and build up of foreign assets,
long run trade balance moderately declines 0.2%, leading to a small change in foreign
assets \(DF\) \((-2\) percentage point of GDP). In other words, the crowding-out, twin-deficit
and human wealth effects are moderate. By comparison, changes in public debt \(DG^s\) are
large thanks to the multiplier effect, which combines increase in fiscal surplus and drop
in insurance premium \((+12\) percentage points of GDP). These changes dominate drops
in other assets, such that net financial assets \(A = DG^s + V + DF\) and total household
wealth \(DG^s + V + DF + H\) increase\(^{27}\). This leads to an increase in welfare \(Q\).

Other experiments confirm the role of the interest rate premium and multiplier effect in
the outcome. When public debt contracts are perfect and there is no premium, public
debt increases due to fiscal reforms lead to welfare losses, consistent with earlier findings
of the literature (Hayakawa and Zak, 2002). This is illustrated by the Perfect contracts
curve in the left bottom graph of figure 5, which plots variations of welfare with labor
income tax increases.

Intertemporal terms-of-trade effects explain the positive spillover impacts on (the for-
eign) country 2. Although twin-deficits leads to an improvement of the trade balance in
country 2, which pushes steady-state foreign debt up, the interest rate increase coming
from crowding-out in country 1 is sufficient to decrease the steady-state level of foreign

\(^{26}\)To remove scale effects, welfare variations \(\Delta Q\) are normalized by the initial output level, \(\Delta Q/Y_0\).
See chapter 1 of the thesis for literature references.

\(^{27}\)One reason for households to be willing to increase financial assets, that is increase savings, relates
to returns. While an increase in savings via firms capital lowers the marginal product of capital and thus
the returns on savings, an increase via government debt does not. Increased demand for capital by the
government, at a high return rate, thus provides an incentive for households to save more.
debt in country 2, dominating other changes in household wealth and welfare in country 2.

Similar mechanisms explain the third outcome, positive welfare impacts of small transfers in the recipient country. Small crowding-out and twin-deficit effects lead to small variations in foreign debt $DF$, firm capital $V$ and human wealth $H$. In comparison, the change in public debt capacity $DG^s$ is large in country 1, thanks to the multiplier effect of improved expected fiscal surplus and declining public debt premium.

The fourth outcome is more unexpected. With large transfers, households’ welfare in the recipient country decline by 0.4% and consumption by 0.6%, while they increase in the giving country, by 0.8% and 1.2% respectively. This outcome is consistent with earlier findings from the literature on transfer paradoxes (Galor and Polemarchakis, 1987; Haaparanta, 1989). Cremers and Sen (2008) illustrate best the intertemporal terms-of-trade effects in the Buiter (1981) model, where the only difference between countries is time preference. If the donor country is more patient, it has a higher propensity to save and accumulates foreign asset positions. The transfer reduces world saving and increases the interest rate. The donor country can then enjoy higher returns on the foreign assets it holds. If the donor is less patient, it has a lower propensity to save and foreign debt. The transfer increases world saving, decreases the interest rate and thus the payment the donor country makes on its foreign debt.

The simulation shows that the transfer paradox result holds numerically when one adds government debt, even when households have the same time preference in the two countries. The fact that small transfers deliver household benefits to the receiving country but not large transfers comes from the convexity of the public debt premium function. As argued in the theory section 3.2, the convexity of the premium function is an intuitive property. Because the premium function is very sensitive for low debt capacity levels, small improvements in the fiscal surplus lead to big declines in the premium and large gains in public debt capacity. The multiplier effect is strong. When the transfer is larger, the larger improvements in the fiscal surplus lead to lower premiums, but in a limited fashion. The premium convexity reduces the strength of the multiplier effect as the transfer and debt capacity increase. The twin-deficit effect becomes stronger, undeterred: the fiscal deficit is reduced and the trade deficit too so the steady-state foreign assets position decline. At some point the multiplier effects is smaller than the twin-deficit effect, public debt gains are dwarfed by foreign asset losses in household wealth, so welfare and consumption decline in the recipient country.

I finish with a detailed discussion of the fifth outcome. Eurobonds decrease the public debt premium from 17.7% to 1.4% in country 1 and increase it from 0.03% to 1.4% in country 2. As a result, the public debt capacity increases from 7% of GDP to 19% in country 1 and decrease from 69% to 60% in country 2. These outcomes are expected. What may be unexpected and paradoxical is the decrease in household welfare in country 1, by close to 2%, and the increase by more than 2% in country 2. There is a redistribution from future generations in country 1 to future generations in country 2.
Why are households in the country which benefits from cheaper financing of public debt suffering from it over the long run?

The analysis of the Eurobond paradox is done in four steps. First, Eurobonds have a double crowding-out effect leading to large trade re-balancing. Assume a shutdown trade channel. The increase in public debt capacity in country 1, due to the declining insurance premium, can only draw from firm capital, pushing the interest rate up. The opposite takes place in country 2. With open trade, households in country 2 are attracted by the higher interest rate in country 1, shift some of their savings to country 1 until interest rates are equalized. Eurobonds thus create a capital flow from country 2 to country 1, re-balancing trade. Country 2 builds foreign assets and country 1 builds foreign debt. Effects are large: the foreign asset position of country 1 declines from 52% to 11% of GDP. The second step concerns the interest rate equilibrium, which declines. Haaparanta (1989) noted that cross-country transfers from low to high saving countries tend to decrease the interest rate. The same happens with Eurobonds. Because country 1 relies more on consumption taxes, it saves more and has a lower autarky interest rate (see section 4.1 for details). Due to trade re-balancing, more of the resources end up in country 1. Taxes are left unchanged and thus do not influence directly equilibrium allocations. The equilibrium interest rate with Eurobonds is thus closer to the autarky level of country 1, lower. The decrease in interest rate stimulates investment and builds capital, the firm value increasing slightly (1% of GDP). Third, the crowding-out effect is associated with a small human wealth effect. Lower interest rate comes with higher wages, increasing labor supply, output and human wealth in both countries. However, effects are small because of the low elasticity of labor supply (human wealth increases less than 4% of GDP). Fourth, the net effect on household wealth differ in both countries. The absolute increase in public debt $DG^s$ is 12% of GDP in country 1, as the convex public debt premium allows for some improvement but only to a limited extent, while firm value $V$ and human wealth $H$ together increase 5%. The trade re-balancing effect on the other hand leads to a sizable foreign asset $DF$ loss of above 40% of GDP. Household wealth $V + DG^s + DF + H$ thus decreases in country 1. So does welfare and consumption. In country 2, intertemporal terms-of-trade effects dominate, as the 40% of GDP increase in foreign assets dwarfs the 9% loss in public debt. Households there gain wealth, welfare and consumption.

To summarize, Eurobonds lead to a wealth, welfare and consumption loss in the recipient country because of the trade effects and the convexity of the public debt premium. The increase in public debt in the recipient country crowds-out capital, has a positive impact on interest rates in that country and thus attract investors from the giving country, leading

---

28Tax increases also lead to an increase in public debt capacity, but trade re-balancing effects are small in magnitude and in the opposite direction. The first reason is that tax increases are domestic policy reforms so the wedge in the interest rates is smaller. The second reason is that tax increases push labor supply down because of the policy change and equilibrium price variations. In contrast, Eurobonds only push labor supply because of equilibrium price variations. With a tax increase, the labor supply reduction dominates the productive capital reduction so that the capital-effective labor ratio increases in country 1.
to trade re-balancing and a sizable loss in foreign assets in the recipient country. By convexity of the public debt premium, increases in the public debt are initially important but quickly reach a ceiling, so that the loss in foreign assets in private wealth dominates the gain in public debt. Conversely, convexity in the premium means that the public debt capacity loss in the giving country is small and dominated by the gains due to trade re-balancing.

To take a real-life example, Germany would pay its public debt higher if it agreed to create a Eurobond to help the finances of the Greek government. But its economic fundamentals are so good and the probability of default so little increased that it would only face a small increase of the debt premium. German commercial banks on the other hand would benefit much from crowding-out and the scarcity of assets in the Greek economy and corresponding increase in returns to private capital.

Additional simulations in appendix D confirm the role of trade re-balancing in the welfare effect of Eurobonds: when the trade channel is shut down so that foreign asset positions are held constant, the Eurobonds paradox disappears: welfare also increases in the recipient country.

4.5 Transition path effects of policy reforms

Policy reform analysis so far only considered long-run effects. Yet, some policies may have short-run effects which dominate long-run effects. In this section I consider the welfare effects along the transition path after the introduction of a Eurobond in country 1, without separating efficiency gains from redistribution between generations and countries.

Figure 2 presents the aggregate welfare level $Q$ in country 1 along the transition path after the introduction of a Eurobond, for different government budget closing instruments. The horizontal line marks the initial welfare level. The dotted line provides the level when the budget is closed with labor income taxes, the plain line with lump sum taxes, the dashed line with consumption taxes and the mixed dash-dot line the closure with government expenditures.

Four observations can be made on the figure. First, aggregate welfare tends to the same value as time grows, whatever the closing instrument. Second, aggregate welfare increases over the short run, except when the budget is closed with government expenditures. Three, short run gains are largest when closing with distortionary instruments, consumption or labor income taxes. Four, short run average welfare gains, if any, are short lived and small in comparison to medium and long run average welfare losses.

Before commenting on each observation, I discuss with some details the transition path variations when the closing instrument is labor income taxes. Figure 3 displays the variation of table 1 variables over time (plus the closing instrument).

On top of the effects which take place over the long run (see section 4.4 for an analysis),
new ones follow from capital adjustment costs and from the temporary tax relief.

The capital stock increases. Because of capital adjustment costs, the increase is smooth. Right after the introduction of the Eurobond, the capital increase is slow. Labor on the other hand adjusts immediately. Initially, the capital-labor ratio is thus low, the interest rate high and the wage rate low, until the smoothing effect of capital adjustment costs wears off, a few periods later.

Effects due to the temporary tax relief last longer. Thanks to the Eurobond, the premium on public debt drops and the country can afford a higher sustainable level of debt. The labor income tax is temporarily reduced, until the debt reaches its new (endogenous) level. Along the transition path, the tax level is adjusted so that the debt follows the arbitrarily given path. The initial tax cut is large, stimulating a large increase in labor supply. The labor supply effect lasts as long as the public debt increases and the tax cut is maintained. As in the comparative statics case, capital is crowded-out by public debt, international capital flows modifies the trade balance\(^{29}\) and the interest rate is lowered. Investment is stimulated, increasing the capital stock. The increase in both production factors leads to large increases in output, allowing for more consumption and higher welfare. Over time, public debt grows, the labor income tax cut becomes smaller, the stimulation of labor and investment declines, the long run effects of Eurobonds become more important until they dominate. In particular, the positive welfare effect disappears over time and becomes negative.

I now provide comments on each of the figure 2 observations, successively. The first outcome is not a surprise, given the simulation approach: closing instruments are needed for transition path outcomes, not for comparative static outcomes (see section 4.2). What-

\(^{29}\)The trade balance initially dips to meet the increased demand in public and private capital, until the long run trade re-balancing effects of Eurobonds dominate.
4. QUANTITATIVE RESULTS

Figure 3: Transition path after Eurobond introduction, labor income tax closure

ever the transition path closing instrument, welfare tends to the same value, given in table 1.

Turning to the second remark, when the closing instrument are lump sum, consumption or labor income taxes, short run effects are similar (and described above). Welfare declines over the short run only when closing with government expenditures. In line with much of the public finance literature, I assumed that government consumption provides no utility to households. The increase in public expenditures, allowed by more affordable and mounting debt, thus provides no utility to households. On the other hand, effects which lead to long run welfare losses start to operate immediately.

The third remark is explained by the distortionary nature of labor income and consumption taxes. The Eurobond introduction allows for higher sustainable public debt, which is built with tax cuts. Reducing taxes which are distortionary has thus a larger welfare impact than lump sum taxes.

The fourth remark is explained by public debt accumulation. The Eurobond allows for a larger sustainable public debt level, reached by tax cuts (or government expenditure increases). These tax cuts are temporary, maintained until public debt has grown to its new steady-state level. Then, tax return to their initial levels and welfare is lower. The positive impact on welfare is thus short lived.

I close this section with a comment on transitions and long-run welfare analysis, build-
ing on the fourth remark. The analysis presented in other sections is appropriate for statements on long-run welfare, not welfare over the transition. In particular, there is no measure of intergenerational redistribution and one cannot isolate pure efficiency gains. The analysis presented in this section shows that there is redistribution from domestic future generations to current generations, but does not separate redistribution impacts from efficiency gains. Figure 2 outcomes however indicate that the introduction of a Eurobond is a net welfare gain over the transition period and the long run only if the social welfare function puts a low, unrealistic, weight on the welfare of future generations. Although the quantitative results of the long-run welfare analysis do not extend to the transition period, qualitative results do, for realistic social welfare functions.

4.6 Multi-dimensional policy reforms

In this section I briefly discuss the concomitant usage of two policy instruments, namely Eurobonds and labor income taxes. I also assess the importance of incomplete debt contracts. Appendix E contains the complete analysis.

Figure 4 provides the main results of the experiment. It shows the debt capacity level $DG^\delta$ as labor income taxes $\tau_w$ vary in country 1, in three cases: when debt contracts are perfect (curve Perf), when they are imperfect (curve Imp) and when they are imperfect and there is a Eurobond (curve EurB).

![Figure 4: Generalized Laffer Curve $DG^\delta(\tau_w)$ for country 1](image)

The figure illustrates a Laffer curve effect: the higher the rate $\tau_w$, the higher the potential fiscal surplus $\mathcal{S}$ and thus debt capacity $DG^\delta$, but the lower the incentive to provide labor supply and thus actual tax revenue.

As expected, the debt capacity is higher under perfect contracts than imperfect contracts, given the absence of a premium and the lower cost of public debt in the first case. What is more surprising is the fact that debt capacity can be almost as high under imperfect...
contracts and a Eurobond as the perfect contract case, provided taxes are raised enough. This property illustrates the cross-country insurance potential of Eurobonds: when a country is hit by a negative public expenditure shock and close to defaulting, the other country may be hit by a positive shock, decreasing the overall likelihood of a default and thus the resulting premium. In the case considered, the potential is large. Other cases, where the no-Eurobond premiums in countries 1 and 2 are close, result in low potential. Just as when tax rates are constant (section 4.4), Eurobonds increase public debt capacity but come with a long-run welfare penalty.

4.7 Summary of quantitative results

Simulations with a strictly convex public debt premium function and small labor supply elasticity show that fiscal consolidation, cross-country transfers and Eurobonds can be successful at raising the public debt level which the economy can sustain. When the gap between public debt premiums in the different countries is sufficiently large, Eurobonds raise the debt capacity level in the high-premium country. While debt capacity effects are similar for all policies, consumption and welfare effects differ. Because they have multiplier effects - fiscal surpluses increases lead to a debt premium decrease - fiscal consolidation and small transfers are welfare improving. Consistent with the transfer paradox literature, large transfers on the other hand lead to welfare losses in the recipient country, because of the premium convexity and intertemporal terms-of-trade effects. The convexity of the premium function leads only to small further increases in debt capacity when transfers become bigger, while the country continues to accumulate foreign debt at a non-deterrable rate. Eurobonds also lead to the paradox that long-run welfare increases in the country which agrees to pay a higher public debt premium and decreases in the other country. In the latter country, increased public debt crowds-out productive capital, creates an interest rate wedge between the two countries and attracts more investments from the foreign country, a trade re-balancing effect which builds foreign debt, depletes household wealth, welfare and consumption.

Transition path analysis show that there are short-run welfare gains after the introduction of Eurobonds if associated with a tax cut. However, the overall welfare improves over the whole transition only if the social welfare function discounts the welfare of future generation at an unrealistically high level.

Simulations in different setups deliver further insights. In certain but not all cases, the cross-country insurance potential of Eurobonds is large. Further simulations, not presented in this paper, allow for instance to quantify the bias from small open economy assumptions or to assess the impact of different specifications for the premium function.
5 Concluding remarks

I have presented an extension of an overlapping generations model to intertemporal trade and incomplete debt contracts with a microfoundation for public debt premiums. I identify conditions under which the public debt premium is convex, a feature consistent with observations of capital markets.

The quantitative analysis shows that policies have large spillover effects, confirming transfer paradoxes exhibited by the theoretical literature (Galor and Polemarchakis, 1987). Eurobonds help improve public debt sustainability but come at a long run welfare cost. They provide short run welfare gains if associated with temporary tax cuts. Fiscal consolidation measures on the other hand can improve both debt sustainability and welfare over the long run, thanks to a multiplier effect.

One limitation of the analysis is the assumption of perfect competition in the financial sector. Distortions from the “too big to fail” type of moral hazard may be larger in that sector than other, so that perfect competition takes place in all but the financial sector. In this case, multiple equilibrium in the government bond markets and self-fulfilling prophecies are possible. The analysis I provide is a benchmark which shows that if there is any welfare value in Eurobonds for high-debt countries, it would come from correction of self-fulfilling prophecies and not from asymmetric shocks to countries. There appears indeed to be no long run welfare gains from cross-country risk sharing via Eurobonds in the absence of self-fulfilling prophecies. Other policy instruments to deal with self-fulfilling prophecies also deserve attention, such as debt restructuring or central banks as last resort lenders (see for instance Grossmann, 2011; Kohler, 2012).

The absence of moral hazard is another limitation of the model. The practical implementation of Eurobonds may indeed come with high-debt countries promises for domestic policy reforms. The cost of the model limitation may however be small. Indeed, the no-moral hazard analysis shows that Eurobonds alone penalizes welfare of high-debt countries while moderate fiscal consolidation improve it, over the long run. Countries which renege on their fiscal consolidation reforms promises would be the ones to suffer.

For robust policy implications over the short run, further research is needed. One area of further research is the extension of the model to imperfect labor markets and multiple sectors. Different policies may have different impacts on unemployment and government budget, via unemployment insurance. Changes in government expenditures also take place via changes in public employment. If one wants to assess the impact of government expenditures changes on the economy, one needs to have several sectors in the model and understand sector re-allocations of labor and the re-allocation costs. The fiscal consolidation exercises analyzed in this paper neglect sector reallocation costs, which may be important over the short run and change conclusions on the welfare impact of consolidation measures.
Appendices
A   Model for quantitative analysis

Four additions are made for more realistic predictions\(^30\). I add lump sum taxes. Capital adjustment costs (Hayashi, 1982) lead to smooth transition paths. The Eurobond premium \(\bar{\pi}\) is the weighted sum of premiums \(P_i\), using debts as weights. I add trade costs, including transportation costs and a premium to cover the risk of private defaults on loans to foreign households, derived from the premium on public debt. The private premium reflects lower foreign investor protection than domestic investor protection and empirical evidence of lower foreign credits during sovereign debt crisis (Arteta and Hale, 2008). Theoretical results hold.

B   Calibration

Other than the specification and parameters for the public debt premium, calibration values are standard in the literature. I choose specifications for the premium function \(\pi\) rather than for the distribution of shocks \(\phi\) (which, via proposition 1 would provide \(\pi\)), for more flexibility. The main specification is strongly convex and derived from Schmitt-Grohé and Uribe (2003). An alternative specification is piece-wise linear, allowing for larger policy shocks.

C   Steady-state equilibrium

Steady-state expressions of optimality conditions in partial and general equilibrium are:

**Lemma 7. (steady-state equilibrium): assuming constant debt over the long run, given an interest rate \(r\), the following hold in each stationary steady-state:**

\[
K = \left( \frac{\alpha}{r + \delta} \right)^{\frac{1}{1-\alpha}} N\varphi_0 \left[ \frac{1 - \tau_w}{1 + \tau_c} \left( 1 - \alpha \right) \right]^{\frac{1}{1-\alpha}} A_0^{\frac{1}{1-\alpha}}
\]

\[
L = \left[ N\varphi_0 \left( \frac{1 - \tau_w}{1 + \tau_c} A_0 (1 - \alpha) K^\alpha \right) \right]^{\frac{1}{1+\alpha}}
\]

\[
DG = -T_1 \left( E \left[ CG_t \right] + Z + \tau_c \left[ \frac{(Y - E \left[ CG_t \right] - Z - I) (1 + \pi_h) - \frac{r(V + H)}{1 + r}}{(1 + \tau_c) \Omega} \right] - T_2 N\varphi \left( \frac{L}{N} \right) - \tau_w w L \right)
\]

\[
C = \left( \frac{- \frac{1}{1+\pi_h}\left(1 + \pi_h\right) (Y - E \left[ CG_t \right] - Z - I) + DG + V + H}{(1 + \tau_c) \Omega - \frac{1+\pi_h}{r}(1 + \pi_h)} \right) + T_2 N\varphi \left( \frac{L}{N} \right)
\]

and steady-state versions of the laws of motion for \(K\), \(V\), \(DF\), \(TB\) and \(TC\) as well as

\(^{30}\)Technical details supporting appendices A, B and C can be found in Davoine (2013).
relationships in lemma 3, where

\[
T_1(DG) = \left( \frac{r + \pi E + \pi(DG/Y)(1 - \mathcal{E})}{1 + r + \pi E + \pi(DG/Y)(1 - \mathcal{E})} \right)^{-1}
\]

\[
T_2 = \frac{(1 + \tau_c)\Omega}{(1 + \tau_c)\Omega - (1 + \pi_h)\frac{1 + \tau_c}{r}}
\]

\[\pi_h = \mathcal{I}_{TB<0}[\xi_0 + id_0 + \pi(DG/Y + \mathcal{R}/Y)]\]

D  Eurobonds and intertemporal trade

To investigate the role of intertemporal trade in welfare effects, I shut down the trade channel by choosing interest rates which maintain the foreign asset position of country 1 (Frozen trade), rather than clearing the goods market (Open trade). Table 2 provides the results.

<table>
<thead>
<tr>
<th>Country</th>
<th>Base</th>
<th>Open trade</th>
<th>Frozen trade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(\mathcal{E} = 1) (\Delta w &gt; 0)</td>
<td>(\mathcal{E} = 1) (\Delta w &gt; 0)</td>
</tr>
<tr>
<td>Interest rate (r)</td>
<td>5.79%</td>
<td>5.73%</td>
<td>5.80%</td>
</tr>
<tr>
<td>Wage rate (w)</td>
<td>1</td>
<td>2.0</td>
<td>+0.2%</td>
</tr>
<tr>
<td>Labor supply (L)</td>
<td>1</td>
<td>33.0</td>
<td>+0.0%</td>
</tr>
<tr>
<td>Capital (K)</td>
<td>1</td>
<td>215</td>
<td>+0.6%</td>
</tr>
<tr>
<td>Output (Y)</td>
<td>1</td>
<td>100.0</td>
<td>+0.2%</td>
</tr>
<tr>
<td>Disp income ((1 - \tau_c)wL + \pi_{lump})</td>
<td>1</td>
<td>56.1</td>
<td>+0.2%</td>
</tr>
<tr>
<td>Trade Balance ((TB - TB_0)/Y_0)</td>
<td>1</td>
<td>-</td>
<td>+1.8%</td>
</tr>
<tr>
<td>Public Premium (\pi(DG/Y))</td>
<td>1</td>
<td>17.7%</td>
<td>1.4%</td>
</tr>
<tr>
<td>Pub Debt Capacity (DG/Y)</td>
<td>1</td>
<td>0.07</td>
<td>0.19</td>
</tr>
<tr>
<td>Firm value (V/Y)</td>
<td>1</td>
<td>2.28</td>
<td>2.29</td>
</tr>
<tr>
<td>Foreign Assets (DF/Y)</td>
<td>1</td>
<td>0.52</td>
<td>0.11</td>
</tr>
<tr>
<td>Assets ((DF + DG + V)/Y)</td>
<td>1</td>
<td>2.87</td>
<td>2.59</td>
</tr>
<tr>
<td>Human wealth (H/Y)</td>
<td>1</td>
<td>5.97</td>
<td>6.00</td>
</tr>
<tr>
<td>Consumption (C)</td>
<td>1</td>
<td>68.8</td>
<td>-2.5%</td>
</tr>
<tr>
<td>Welfare ((Q - Q_0)/Y_0)</td>
<td>1</td>
<td>-</td>
<td>-1.7%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-</td>
<td>+2.3%</td>
</tr>
</tbody>
</table>

*Open trade*: general equilibrium (from table 1). *Frozen trade*: interest rate for constant foreign assets \(DF_1\).

Table 2: Eurobonds and fiscal reform with open or frozen trade, long run effects

The two main observations are that Eurobonds become welfare increasing when the trade channel is shut down and that there are barely any differences for labor income tax changes.

The first observation illustrates the importance of trade re-balancing in the Eurobonds paradox, the fact that households in country 1 lose welfare, in the general equilibrium case. Firm value \(V\) and human wealth \(H\) change little, while the convex public premium leads to a significant but moderate increase of public debt \(DG\). The demand on capital (crowding-out) pushes interest rates up in country 1, attracting inflows of savings from
country 2. This trade re-balancing drops much long run foreign assets $DF$, dominating the gain in $DG$. Private wealth $V + H + DG + DF$ and welfare $Q$ drop in country 1.

When the trade channel is shut down, foreign assets $DF$ are unchanged and the gain in public debt $DG$ is the dominating change in private wealth. Welfare then also gain.

In contrast, fiscal consolidations are not affected by shutting down of the trade channel. Due to the convexity of the debt premium and the multiplier effect, small changes in taxes have a big impact, increasing much the public debt capacity. Small tax changes being sufficient, labor supply incentives and interest rate are little affected. In the open trade case thus, households have little incentives to ship savings in the foreign country. There is essentially no trade re-balancing under open trade. Shutting down the channel thus leads to similar outcomes.

### E Multi-dimensional policy reforms analysis

In this section, I consider joint policy reforms in different legal environments, comparing the effect of labor income tax variations in three cases: perfect debt contracts, imperfect debt contracts and imperfect debt contracts with a Eurobond in place. There are still two countries with free trade and general equilibrium. I change the calibration to get a more explicit illustration, using a higher elasticity of labor supply and a piece-wise linear function for the public debt premium, which tolerates larger policy shocks.$^{31}$

I compare long run outcomes over an entire range of tax values, computing the equivalent of Laffer curves for every variable of interest. The higher the tax $\tau_w$, the higher the potential fiscal surplus $\mathcal{S} = \tau_c C + \tau_w wL - CG - Z$ and debt capacity $DG = \sum_u \mathcal{S}_u / (R_u + \pi(DG / Y_u))$ but the lower the incentive to provide labor supply. Past a certain level of tax, labor supply will be so low that debt capacity will decrease with further increases in taxes. I will call the curve $DG(\tau_w)$ the generalized Laffer curve.

The main results of the experiment are presented in the top left graph in figure 5, which shows the generalized Laffer curve $DG(\tau_w)$ in country 1 for the three cases: perfect debt contracts (dashed line), imperfect debt contracts without Eurobonds (dotted line) and imperfect debt contracts with Eurobonds (plain line).

The graph shows three interesting outcomes. First, the public debt capacity level is higher under perfect debt contracts than under imperfect contracts. Second, the generalized Laffer curves have a hump shape. Third, Eurobonds provide almost perfect insurance, in the sense that the maximum debt capacity level over all tax values under imperfect contracts and Eurobonds is almost as high as the maximum level under perfect contracts. While the first two outcomes are no surprise, the third outcome may be unexpected.

---

$^{31}$Rather than a value of 0.2 for the labor supply elasticity, consistent with microempirical estimates, I take a value of 2, closer to the value 4 used in several macroeconomic studies (for instance, Prescott, 2004).
Figure 5: Variations with the labor income tax rate $\tau_w$ in country 1
The fact that the debt capacity level is higher under perfect contracts than imperfect contracts is due to the risk of government defaults. With perfect contracts, investors are fully protected against defaults and do not need to charge any insurance premium \(\pi(DG^u)\), so the debt capacity level \(DG^u = \sum_u \mathcal{S}_u / (R_u + \pi(DG^u/Y_u)) = \sum_u \mathcal{S}_u / R_u\) is higher.

The hump shape of the generalized Laffer curves follows from the hump shape of the traditional Laffer curve, recording government labor income tax revenue as the tax rate is increased and illustrated in the second graph on the top of figure 5. The hump is due to the disincentive effects of taxes on household labor supply, shown on the left third row graph.

The surprise in the third outcome is how close Eurobonds are to perfect insurance; in other words, how close the maximum debt capacity level is from the case with perfect debt contracts as taxes vary.

In the perfect debt contract case, investors are able to recover assets they loan to government whatever the path of shocks to government expenditures. These loans allow government to provide public goods whatever the shocks and the need to temporarily increase the amount of public goods they provide. Investors are not afraid of losing assets (government default). When public debt is high, due to a series of large negative shocks on government expenditures, investors know that the government will sooner or later face a series of small shocks on expenditures, which will reduce public debt and allow debt servicing. Investors loans thus provide perfect insurance to governments.

In imperfect debt contract cases, investors put a limit to the amount they are willing to lend to governments. When public debt is high after a series of large negative shocks on expenditure, they ignore (or are forced to ignore by the shallowness of their pockets) the statistical nature of shocks and the fact that shocks will eventually be small. Governments can only insure against negative shocks to their expenditure up to a point. The insurance is imperfect.

With a Eurobond, the low-premium country agrees to carry some of the default risk from the high-premium country. A Eurobond thus has a cross-section insurance dimension. When one of the country is hit by a negative shock and is closer to defaulting, the other may be hit by a positive shock, decreasing the overall likelihood of a default. When shocks are uncorrelated, the investor is thus less likely to suffer from a default from the two countries at the same time, allowing him to lower the insurance premium.

In contrast, the imperfect debt contracts case corresponds to pure self-insurance along a time-series dimension. The high-premium country does not benefit from the low risk of default from the low-premium country and needs to pay a high premium.

There is however a price to pay for higher public debt capacity with Eurobonds. As shown in the bottom left graph of figure 5, household welfare is always lower with Eurobonds.

The explanation is identical to the case of the introduction of a Eurobond, presented in
section 4.4, with some adjustments due to the differences in calibration. The explanation follows the same four steps. First, Eurobonds have a trade re-balancing effect. The rise in public debt capacity in country 1 crowds-out productive capital and push the interest rate higher in country 1, while the opposite takes place in country 2. Households in country 2 ship savings to country 1 to maximize capital returns, a trade re-balancing flow. Second, the equilibrium interest rate drops. As country 1 relies more on consumption taxes and saves more, its autarky interest rate is lower. With trade re-balancing, more of the resources end up in country 1. The equilibrium interest rate with Eurobonds is thus closer to the autarky level of country 1, lower. Third, the lower interest rate comes with higher wages, so households in both countries provide more labor supply, when tax rates are still low. Output follows and increases as well. Because the elasticity of labor supply is large, effects are sizable: output increases 7% in country 1 and 11% in country 2 for low tax rates. Fourth, increased output allows for more consumption in both countries for low tax rates. The welfare impact however differs. Which of the increase in consumption utility and labor disutility dominates depends. Trade re-balancing comes with a higher foreign asset position in country 2 and lower position in country 1. It has a positive impact on private wealth in country 2 and a negative impact in country 1. Since welfare is nothing but the propensity to consume out of private wealth, welfare improves in country 2 and declines in country 1.

I finish this section with two remarks. First, welfare is higher in the imperfect contract case without Eurobond than in the perfect contract case, surprisingly. The explanation follows that of the Eurobond paradox: in the perfect contract case, higher public debt crowds-out private capital, pushes interest rates up in country 1, attracting country 2 investors, depleting country 1 foreign assets over the long run, reducing wealth and welfare. Second, the possibility of Eurobonds to lift public debt capacity in high-premium countries depends on the premium paid in other countries. Simulations where country 2 has higher public expenditures, increasing the premium from 1% to 24% (compared to 29% in country 1) show that Eurobonds do not help, even as taxes increase.
Bibliography

debt service: Some estimates for OECD countries. *The B.E. Journal of Macroecon-
nomics* 7(1), 28.


*Journal of International Economics* 74(1), 53–69.

new approach to understanding the effects of fiscal policy on saving. *Scandinavian

EU structural funds on regional performance. *Journal of Public Economics* 94(9-10),
578–590.

Economy* 93(2), 223–47.


Curriculum Vitae

Education

2008-2013 PhD in Economics and Finance, University of St. Gallen
1999-2000 Master of Science in Operations Research, Rutgers University, USA

Employment

2012-... Applied research, Institute for Advanced Studies, Vienna, Austria
2009-2011 Research and teaching assistance, University of St. Gallen
2008-2009 Management consulting, KPMG, Genève
2005-2008 Management consulting, PricewaterhouseCoopers, Genève
2001-2005 Management consulting, Eurogroup, Paris, France
1998 Research, Universidad Nacional Autonoma de Mexico, Mexico

Thomas Davoine, December 2013