How Should Capital Be Taxed?*

Spencer Bastani† and Daniel Waldenström‡

September 18, 2019

Abstract:
This paper analyzes the theoretical and empirical literature on capital taxation in advanced economies, highlighting the latest developments in the optimal tax literature that have challenged some of the conventional wisdoms. We review and discuss the main arguments against and in favor of capital taxation and what they have to say about specific capital taxes on capital income, wealth, property, inheritance and corporate profits. We also describe stylized empirical facts on the current practices of capital taxation in the Western world, empirical evidence on the distribution and aggregate importance of wealth and capital income, and make an assessment of the distortionary effects of different capital taxes. Our overall conclusion is that capital taxation has an important role to play in an optimal tax system, but when considering specific capital taxes, we find that only some of these are able to strike a balance between an optimal and administratively feasible implementation.

Keywords: Optimal taxation, Capital taxation, Wealth tax, Inheritance tax, Income inequality, Wealth inequality.


* This is a substantially re-worked version of Bastani and Waldenström (2018) which circulated under the title “How Should Capital Be Taxed? A Swedish Experience”. We have received valuable comments from colleagues and seminar participants at the OECD and the IIPF annual conference 2018 as well as Annette Alstadsæter, Bas Jacobs, Sebastian Köhne, Etienne Lehmann, Henrik Kleven, Thomas Piketty, Emmanuel Saez, Guttorm Schjelderup and Floris Zoutman. Financial support from the Swedish Research Council and the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged.
† Department of Economics and Statistics, Linnaeus University, Sweden; CESifo, spencer.bastani@lnu.se.
‡ Research Institute of Industrial Economics (IFN), Sweden; CEPR, CESifo, IZA, daniel.waldenstrom@ifn.se.
1. Introduction

The taxation of capital is a question that has spurred substantial academic and political dispute for a long time. Perhaps this is not that surprising given the complex nature of capital. Capital is needed to fund investments, it provides consumption benefits, as in the case of housing wealth, and it serves as a vehicle for individuals to transfer resources across time, jurisdictions and generations. Empirical studies have shown that capital is more unevenly distributed than labor income, and also that the aggregate value of the capital stock relative to total labor income has grown in wealthy nations in recent years. Despite all of this, it is fair to say that a substantial uncertainty still remains regarding the economic effects of capital taxes and their role in the tax system.

The purpose of this paper is to review the current theoretical and empirical literature on capital taxation in order to allow us to address the question of how personal taxes on wealth and capital income should be designed in advanced economies. We discuss the current theoretical research on the subject of capital taxation in an attempt to provide a unified discussion about the optimal taxation of capital with an eye towards practical policy recommendations. The survey complements earlier articles on the connection between optimal tax theory and tax policy, such as Mankiw, Weinzierl and Yagan (2009), Banks and Diamond (2010), Diamond and Saez (2011) and Jacobs (2013).

The theoretical discussion starts in the modern optimal tax literature, and is divided into two parts. First, we analyze arguments against and in favor of the taxation of capital income. Among the arguments against such taxes are those focusing on their distortions of the inter-temporal allocation of consumption and physical investment. Arguments in favor of capital taxation center around aspects of both equity, such as the fact that individuals are heterogeneous and that taxing capital income enables more redistribution than if labor income would be taxed alone, and efficiency, such as the taxation of economic rents, reducing distortions on human capital accumulation and mitigating the distortions associated with progressive labor income taxation. The second part discusses theoretical considerations concerning specific capital tax bases: wealth taxation, property taxation, personal and corporate income taxation and inheritance taxation.

We also provide an empirical discussion, portraying the structure of capital taxation in today’s developed world, and then continue with an analysis of the aggregate importance and distribution of wealth and capital income, the implications of hidden offshore wealth and the recent appearance of information exchange agreements across countries, and finally what is known about behavioral responses to capital taxation. It should be noted that questions about the distribution of capital have not played a prominent role in traditional theories of optimal capital taxation, but recent studies highlight its importance.

The paper is organized as follows. Section 2 begins by describing the aggregate importance of wealth and capital income. Section 3 reviews the theoretical literature on optimal capital taxation and the main
arguments against and in favor of the taxation of capital. Section 4 discusses how capital is taxed in practice, offering cross-country comparisons, and examines the literature on behavioral responses to taxes on wealth and capital income, as well as offering international perspectives on capital taxation, with a focus on the role of hidden offshore wealth and information exchange agreements. Finally, section 5 offers a concluding discussion, delivering a number of concrete policy recommendations, and discusses limitations and ideas for future research.

2. Capital taxation and wealth distribution: An international outlook
This section describes a set of empirical patterns that provide the background for our analysis of capital taxation. First, we discuss capital taxation practices in today’s developed economies, including their importance relative to other taxes and how they are structured across tax bases. Second, we examine the aggregate size of private capital and the distributional characteristics of personal wealth and capital income in rich countries.

2.1 Current practices of capital taxation
A capital tax refers to any tax on the return to savings, capital gains, dividend income, firms’ profits (corporate taxation), property taxation, inheritance/estate taxation and wealth taxation. Sometimes it is useful to divide these taxes into two categories depending on whether or not the tax is levied on an income stream (flow taxation) or on the stock of capital. In some cases, it does not matter from an economic perspective whether the stock or the flow is taxed. For example, if the annual rate of return on an investment is 4 percent, an annual wealth tax of 1 percent is equivalent to a capital income tax of 25 percent (in terms of the total annual tax burden). In the canonical life-cycle model of labor supply and savings, taxes on the stock and the flow are equivalent. However, as will be discussed below, if rates of return differ across individuals, or are uncertain, taxes on the stock of capital and capital income taxes are no longer equivalent. Moreover, wealth is a broader concept than income, which includes assets such as pensions and closely-held shares, which further break this equivalence from a conceptual point of view.

Capital taxation consists of a number of taxes on different forms of capital-related income flows, stock values and transfers. One way to assess the fiscal importance of capital taxation is to look at the aggregate revenues of capital taxes as a percentage of GDP. Figure 1 shows these outcomes for a selection of OECD countries in 2016. Panel (a) shows the overall tax-to-GDP and its main components, and from these numbers it stands clear that capital taxes represent a relatively small share of total tax receipts. Their average share is about one tenth of all tax revenues and there is not much variation across countries. Direct labor taxes constitute the largest source of tax revenues followed by indirect consumption taxes.
Is it surprising that capital taxes are of such relatively limited fiscal importance? Perhaps not. The numbers make sense if we relate them to macroeconomic fundamentals. For example, assuming a private wealth-to-GDP ratio of 500 percent (the ratio ranges between 200 and 600 percent in developed economies), a real rate of return to capital of three percent and a capital income tax of 30 percent, this would result in capital tax revenue at 4.5 percent of GDP (500% × 3% × 30%). Alternatively, if one uses the net capital share of national income, which is at around 20 percent (ranging between 10 and 30 percent in rich countries), and a capital income tax of 30 percent, this would result in a capital tax revenue of 6 percent of GDP (20% × 30% = 6%).

Panel (b) in Figure 1 shows the composition of different capital tax bases for the same set of countries. The corporate tax is the most important capital tax, representing roughly half of total capital tax revenues. Property taxes, which include both recurrent taxes (on either the stock or the value of imputed property income) and one-time transfer taxes, are also relatively important, representing on average one third of all capital taxes. The other capital taxes are smaller and include taxes on households’ interest and dividend income, realized capital gains and, in a few countries, taxes on net wealth and intergenerational transfers.

Figure 1: Tax revenues in OECD countries (percent of GDP).

Information exchange agreements between countries with the purpose to obstruct tax evasion and tax-driven capital flight to tax havens are important parts of capital tax structures. Since the mid-2000s, cross-national organizations such as OECD, G20 and the EU have initiated the signing of bilateral and multilateral agreements among both developed and developing countries. These efforts largely consist of introducing reporting standards and automated information exchange arrangements, aimed at curbing
tax evasion and tax planning.\textsuperscript{1} The number of participating countries have increased steadily, from zero in the early 2000s to 129 in late 2019.

The effectiveness of information exchange agreements rely on the number of participating countries and the credibility of the sanctions set up by the participating countries. In a survey of the impact of the early agreements, OECD (2017) estimates that over 500,000 taxpayers have disclosed assets over the past eight years, resulting in an increase of over 85 billion euros in tax revenues. Johannesen and Zucman (2014) found evidence on tax evaders moving to tax havens outside the treaties, thus suggesting that the agreements indeed seem to have had an effect on investors. Despite these indications, much uncertainty remains concerning the total impact of the agreements, not least concerning the corporate sector where dimensions of legality are less easily defined.

\subsection*{2.2 Distributional evidence}

In Figure 2, we present two general trends in the Western world: a) the ratio of total aggregate private wealth over national income and b) the share of wealth held by the richest 1 percent. Looking at the aggregate wealth-income ratio, the evidence shows that it has grown from approximately 200-300 percent (two-three years of national income) in the 1970s and the 1980s to a level twice as high today, approximately 400-600 percent of national income.\textsuperscript{2} Many different factors account for this upward trend, including the deregulation of postwar markets, relative asset price gains (especially in housing markets) and the marketization of funded pension assets.

The level and trend of wealth inequality matters for capital taxation through equity concerns. Estimating wealth inequality is however more difficult than estimating income inequality due to problems with the definition, valuation and measurement of private assets and liabilities. Despite these difficulties, panel b of Figure 2 shows some recent attempts to estimate the top 1 percent wealth share since 1970 using consistently defined data for a handful of Western economies. The results indicate a scattered overall picture: Denmark, France and the UK have experienced an almost stable top percentile wealth share, around 20-25 percent, during the entire period (the UK is down from almost 30 percent in 1970) whereas the US seem to have experienced increased wealth concentration, although the trend (and level) depends quite starkly on which data source (survey or capitalized income data) or method (of capitalizing incomes) one uses.\textsuperscript{3}

\footnote{Among these initiatives are OECD’s Base Erosion and Profit Shifting (BEPS), EU’s Anti-Tax Avoidance Package (ATAP) and the US’s Foreign Account Tax Compliance Act (FATCA).}

\footnote{See, for example, Piketty and Zucman (2014, 2015) and for Sweden also Waldenström (2016, 2017). Note that we define capital as the sum of asset values less the value of liabilities, and thereby use the terms capital and wealth interchangeably.}

\footnote{The discussion about the US wealth inequality levels and trends is ongoing and it is fair to say that consensus is not reached. The traditional wealth data comes from the Survey of Consumer Finances and they indicate a fairly stable top percentile wealth share (see, Wolff 2017, Bricker, Henriques and Hansen 2018). Saez and Zucman
Figure 2: Wealth-income ratios and Top 1% wealth shares

Notes and Sources: Panel (a) shows the ratio of aggregate private wealth (sum of all assets less liabilities of households and non-profit institutions serving households) to national income. Data come from www.wid.world for Denmark, Piketty and Zucman (2014) for France, UK and USA, and Waldenström (2017) for Sweden. Panel (b) shows the share of the same private wealth as in (a) that is held by the richest percentile of the distribution of adult residents. Sources are Jakobsen et al. (2019) for Denmark, Garbinti, Goupille-Lebret and Piketty (2019) for France, Wolff (2017) for “USA (SCF)”, Saez and Zucman (2016) for “USA (SZ)” and Smith, Zidar and Zwick (2019) for “USA (SZZ)”.

Hidden wealth in tax havens could potentially alter the picture of the aggregate size and distribution of wealth and capital income. Despite a sizeable literature on tax-driven capital flight, the complex nature of the data prevents strong conclusions concerning the links between international capital mobility, wealth distribution and tax evasion. Most studies find that there is evasion and that the amount evaded is significant. Zucman (2013) attempted to estimate the extent of hidden offshore wealth globally using an ingenious approach based on netting out financial assets and liabilities in national balance sheets with the purpose of identifying unexplained gaps. His finding was that approximately USD 6 trillion, or 8 percent of global wealth, was placed in tax havens in 2007. Statistics from the Swiss tax authorities presented in Johannesen (2014) suggest that as much as 80 per cent of all wealth that Europeans placed in Switzerland is not reported in their respective countries. Annual tax losses due to tax evasion are also significant, estimated to range between 300 and 1000 billion globally, of which the majority of these (2016) used capitalized capital income data and found a significant increased top wealth share level throughout the 1980s-2000s, but their estimates have been questioned by several papers (Kopczuk 2015, Bricker et al. 2018, Smith, Zidar and Zwick 2019).
are concentrated to OECD countries (Crivelli, De Mooij and Keen 2016). The distributional consequences of this tax evasion are also difficult to estimate and the studies attempting to do so find that the results are sensitive to assumptions.4

Wealth mobility and intergenerational transfers are other inequality dimensions with relevance to the analysis of capital taxation. Little is known about these outcomes in most countries because of the unusually high data requirements. Studies of different countries have shown that intergenerational wealth mobility tends to be lower than intergenerational income mobility (see, for example, Charles and Hurst 1996 for the US, Adermon, Lindahl and Waldenström 2018 for Sweden; Boserup et al., 2017, for Denmark). The role of inheritances appears to be of specific importance: Adermon et al find that that a large part, perhaps half, of the expected wealth rank transmission from one generation to the next adheres to inheritance and gifts.5 A similarly large role of intergenerational transfers, but for intergenerational income transfers in the top of the income distribution, was found by Björklund, Roine and Waldenström (2012). They found that the role of family background was stronger when capital income is considered alongside labor earnings, and that this pattern seems largely driven by transfers. Inherited wealth flows at the aggregate level have been studied in some countries, relating these flows to national income or the stock of total wealth. Findings for France (Piketty 2011) and Sweden (Ohlsson, Roine and Waldenström 2019) suggest that inherited wealth flows as share of total income have become increasingly important in the recent decades.6

3. Theory

To approach the question of the most desirable way to tax capital, a framework for the analysis is needed that specifies the objective of tax policy as well as the relevant constraints facing the policymaker. Most of our discussion will be based on the modern approach to optimal taxation, initiated by Mirrlees (1971), where the government balances the gains from redistribution and the financing of several private and public goods, with the harmful effects of taxes on economic activity.

The starting point of optimal tax theory is the study of a population of taxpayers that differ in terms of their skills, or capacities, to earn income.7 If the government could observe each individual’s skill level,

---

4 For example, Alstadsæter, Johannesen and Zucman (2019) are able to link some of the named individuals in the renowned “Swiss leaks” and “Panama papers” to Nordic administrative registers and show that they belong to the top percentile of the wealth distribution. Roine and Waldenström (2008, 2009) examine the case of Sweden and estimate the impact of hidden wealth on top income and top wealth shares, finding that both shares increase significantly although not enough to change the overall inequality trends in recent years.

5 This finding is in line with a study by Boserup, Kopczuk and Kreiner (2018) who documents that intergenerational wealth correlations are higher for Danish children and young adults with deceased parents or grandparents.

6 The share of wealth that is inherited amounts share varies between 30 and 60 percent in Western countries (Wolff 2015; Piketty and Zucman 2015, Ohlsson et al. 2019).

7 Throughout this text, we will use the terms “skills”, “abilities” and “capacities to earn income” interchangeably.
the tax planner could assign each individual a tax or transfer depending on their unique personal capacity to earn income. Such a hypothetical tax system would fulfill all of society’s distributional objectives, whatever those may be, without disrupting economic activity. Individual economic circumstances are not, however, observable to the government and individuals have no incentives to truthfully reveal them. For this reason, taxes must be based on observable characteristics and economic quantities, such as income or wealth. This causes economic distortions as individuals will change their income and wealth in response to taxation. It should be emphasized, that the economic costs of taxation (often referred to as excess burden or deadweight loss) are rooted in the information problem of the government, namely, the government’s inability to verify individual earning ability. As we will explain in more detail below, capital taxes can serve an important function by mitigating the information asymmetry between the government and private agents.

Formally, an optimal income tax problem amounts to the maximization of a social welfare function which describes how the welfare of agents in the economy should be valued and aggregated. This maximization is subject to a set of incentive constraints and a government’s budget constraint. The incentive constraints capture that individuals freely choose their desired income subject to the taxes set by the government. This implies that whenever the government attempts to raise the welfare of low-skill individuals through changes in taxes and transfers, it has to fear that high-skill individuals might adjust their income in an attempt to replicate the income of low-skill individuals in order to qualify for a lower tax burden. The social welfare function is typically formulated as a function of individual utilities. One such social welfare function is the ‘Utilitarian’ one where the government is maximizing the sum of individual utilities. The goal of redistributive taxation is then to equalize marginal utilities across individuals. Another social welfare function is the so-called max-min social welfare function, where the government’s aim is to maximize the utility of the lowest skilled agent in the economy. If the lowest skilled agent in the economy does not work, the max-min social objective is equivalent to the objective of tax revenue maximization from the working population. Notice that the social welfare function embodies a normative assumption, and is specified by the researcher. Researchers often compute optimal policies under different social welfare functions with the hope of identifying desirable features of tax systems that are fairly robust with respect to these assumptions.

The original Mirrlees (1971) model was static. However, subsequent contributions have also analyzed richer, multi-period, Mirrleesian economies. In such models, researchers need to specify how the labor productivity of agents evolves over time. A common approach is to view individual productivities as partially pre-determined (depending for example on inherited traits, the childhood environment, access

---

8 This is just one way of describing that there is a positive labor supply elasticity.
9 See Bastani and Lundberg (2017) for a comprehensive analysis of the implicit social welfare weights inherent in Swedish labor income taxation 1971–2012, assuming taxes have been set optimally.
to education etc.), partially evolving over time (as a consequence of circumstances, such as luck, and health conditions), and partially being the result of economic choices (such as the investment in education, on-the-job training etc.). This implies that, at any point in time, the distribution of utility depends on the initial heterogeneity in the economy (what individuals are born with), current and past realizations of economic shocks as well as individuals’ past economic choices.\(^\text{10}\)

A persistent feature of the economy is that capital income is more unevenly distributed than labor income. Thus, from a purely distributional point of view, capital taxes appear desirable as instruments to combat inequality. The relevant question, however, is to which extent capital taxation enables the tax system as a whole to more efficiently raise tax revenue and achieve distributional objectives. Our discussion about the optimal taxation of capital will focus mainly on the desirability of capital taxation in economies where labor income is already subject to progressive income taxation. The relevant question then becomes whether or not taxing capital income in addition to labor income allows for more income redistribution than what can be achieved with a non-linear tax on labor income alone. Moreover, the fundamental task is to balance the equity gains from taxing capital income against the distortions of capital income taxation, in the form of distorted savings decisions, labor supply distortions and distortions in portfolio choice.

In the early macroeconomic models used to study optimal capital taxation, the analysis centered on the dynamic decisions of a representative individual, focusing mostly on the efficiency properties of a tax system that raises a given amount of revenue. In these models, distributional concerns were absent. At the same time, introducing heterogeneity in terms of skills, as in the Mirrlees (1971) framework, did not appear to change the result that capital income should not be taxed, at least not in the simple setting of Atkinson and Stiglitz (1976).

Today, these models serve as important theoretical benchmarks. If all inequality in capital income originates from inequality in labor income (because of differences in work ability), it is perhaps not surprising to find an unimportant role for capital income taxation in the optimal tax system. The opposite extreme would be a situation where all inequality derived from inequality in capital income. In such a situation, the only way to achieve redistribution would be to tax capital income.

One major development in the recent research literature is that researchers now are beginning to explore the implications of individual heterogeneity beyond differences in labor market ability for the design of

---

\(^\text{10}\) The underlying reasons for why skills differ, matter for the interpretation of government interventions to reduce inequality. The extent progressive income taxation can be regarded as redistribution and the extent it can be regarded as insurance particularly carry weight in political discussions about the design of the tax system. In practice, making a distinction between redistribution and social insurance is difficult as it is hard for an empirical researcher to assess whether the inequality in outcomes that is observed in the data is the result of choices, predetermined characteristics, or chance (good or bad luck).
optimal tax systems.\textsuperscript{11} The most attractive reason to tax capital income, in our view, is the regular empirical finding that there is substantial heterogeneity in capital income conditional on labor income.\textsuperscript{12} Thus, taxing capital income in addition to labor income allows to achieve more redistribution than taxing labor income alone. The normative implications depend on where the inequality in capital income conditional on labor income derives from. The literature has recently highlighted, for example, heterogeneity in bequest behavior, in the likelihood to receive and give bequests, and in investment returns. However, all these heterogeneities can lead to a role for positive optimal capital income taxation.\textsuperscript{13}

Notice that the correlation between labor income and capital income, and the extent to which the labor income tax is allowed to be progressive, plays a crucial role in understanding the redistributive role of capital taxation. If labor income perfectly predicts capital income, then there is no equity reason to tax capital provided that labor income can be taxed in a flexible nonlinear way. To the extent that the correlation is imperfect, complementing labor income taxes with capital income taxes enables to redistribute between people with different capital income who share similar labor income. However, if there are restrictions on the extent to which the labor income tax is allowed to be nonlinear (for instance if the labor income tax is restricted to be proportional), then capital taxes are useful for redistributive purposes even if the correlation between labor income and capital income is perfect and there is no heterogeneity in capital income conditional on labor income. In this case, capital income taxes serve to achieve the progressivity that is missing in the labor income tax.\textsuperscript{14}

Another major development is that the literature has begun to connect theories of optimal capital taxation to the distribution of capital and the elasticity of capital supply with respect to the after-tax return (in terms of sufficient statistics). Saez and Stantcheva (2018) provide a framework in which many policy questions about capital taxation can be addressed, including the role of heterogeneous returns and differences in preferences for different types of wealth.

A limitation of many modern studies of capital taxation is that they are set in a closed economy where the efficiency costs of capital taxation relate to how individuals change their intertemporal consumption patterns and how capital taxation discourages productive domestic investments and growth. In an open

\textsuperscript{11} The fact that a vast majority of studies of optimal capital taxation consider models with a single dimensional heterogeneity is not because scholars consider this to be the most appropriate assumption, but rather that there are severe mathematical difficulties involved in solving for optimal income taxes in economies with multidimensional taxpayer heterogeneity. The literature on optimal nonlinear taxation in the presence of multidimensional heterogeneity and multiple tax bases has been greatly advanced in recent years; see Golosov et al. (2014) and Lehmann et al. (2018). See also Hermle and Peichl (2018) who analyze optimal linear taxation of multiple tax bases.

\textsuperscript{12} See Bastani and Waldenström (2019).

\textsuperscript{13} Another type of heterogeneity stems from life-cycle considerations in overlapping generations models, as individuals in different ages have different capital income (even conditional on having the same labor income). In these models, capital income taxation becomes a substitute for age dependent taxation. We discuss this below.

\textsuperscript{14} The correlation between labor income and capital income plays an important role when constructing practical policy proposals. In particular, if one wishes to lower marginal labor income taxes for efficiency reasons, raising capital income taxes can be an important tool to off-set the distributional implications.
economy, additional efficiency costs arise to the extent that individuals and can migrate and/or move their wealth abroad.

Most countries’ tax systems abide by the so-called residence principle, which means that individuals are liable to pay taxes on all their incomes, independently of where these incomes have been earned. An important determinant of the economic costs of capital taxation is the possibilities for individuals to engage in tax evasion and tax planning, thereby avoiding taxation in their home country. As there is a clear upward trend in terms of information exchange agreements between countries, the possibilities to avoid taxation in the home country are diminishing. This increases the capacity of small open economies to tax capital. If all tax planning and tax avoidance possibilities disappear, the only way for an individual to avoid taxation in the home country is to migrate.

3.1 Why capital income should not be taxed
One way to approach the issue of capital taxation is to study a neoclassical growth model where an infinitely lived representative individual supplies labor supply in each time period and transfers resources across time periods through savings in order to smooth consumption. The savings of the representative individual finances the investments in the economy and the optimal tax problem is to design taxes on labor and capital income in every time period in order to reach a given amount of tax revenue in the most efficient way (maximizing the welfare of the representative individual). In such a framework, Chamley (1986) argued that the tax on capital should be zero in the long run. The same conclusion was reached by Judd (1985) who studied redistribution from capitalists to workers in a closely related framework.

These studies have been extremely influential and have established a widespread intuition that capital income taxation is undesirable because it becomes very distortive over long time horizons. The reason being that the distortion from capital income taxation grows over time as the interest on savings becomes increasingly important to finance future consumption the further one looks into the future. Many scholars have questioned the Chamley-Judd result because it rests on unrealistic assumptions. For example, infinite planning horizons reflect the decisions applying for a dynasty, where different generations are perfectly connected through altruistic bequests, and neglects the inequality that is created over time between individuals who receive and individuals who do not receive inheritances. We will come back to these issues below, but the perhaps most important criticism against the result is that it might be logically incorrect. Straub and Werning (forthcoming) argue that the steady state optimal (long-run) capital tax in fact can be very large (equal to an upper bound) due to the possibility that income effects

15 The infinite horizon of the representative individual can be interpreted as an infinite dynasty where different generations are perfectly linked through inheritance.
dominate substitution effects, implying agents save more when they are being taxed more, leading to increasing benefits of raising capital taxes over time.\textsuperscript{16}

A different starting point for analyzing capital income taxation are models building on Mirrlees (1971). In this framework, the government desires to redistribute between individuals with different abilities to generate labor income, but since ability is assumed to be private information has to do so indi\textit{rectly} by taxing observable economic quantities (such as labor and capital income). A relevant question that arises is whether enhanced redistribution can be achieved by taxing capital income in an economy where labor income is already subject to progressive taxation.

The relevant starting point to think about this issue is Atkinson and Stiglitz (1976), one of the most influential studies in public finance. Atkinson and Stiglitz provided conditions under which it is undesirable to use differentiated commodity taxation if the government is allowed to simultaneously use nonlinear income taxation to pursue redistributive goals. More specifically, they showed that if preferences are homogenous across individuals and weakly separable between consumption and leisure, all information about individuals’ earnings abilities is contained in their choices of labor income, and differentiated commodity taxation would only create distortions without resolving any of the asymmetric information between the government and private agents. Since consumption in different time periods can be viewed as different commodities, and a tax on capital income is equivalent to making consumption tomorrow more expensive than consumption today, their work has been interpreted to imply that capital income should not be taxed. Today, however, many scholars view the application of the Atkinson-Stiglitz result to capital taxation to be an unfortunate misinterpretation and oversimplification of the issue of capital taxation, see Stiglitz (2018).

While intuitively appealing, the result is not robust to perturbations in the modelling framework.

First of all, Atkinson and Stiglitz analyzed a model where individuals live for two time periods, and work only during the first (later studies have extended the analysis to life spans over several periods). A zero tax on capital income is then optimal only if the labor income tax is allowed to be a complicated function of current annual income, as well as annual income each year in the past. Such labor income taxation does not exist in practice, restricting the policy relevance of the application of the Atkinson-Stiglitz result to the issue of optimal capital income taxation.

\textsuperscript{16}See also Jacobs and Rusu (2017). The main message of Chamley-Judd should however not be completely dismissed, since even for modest planning horizons, the compounding effect of capital income taxation can become quite strong. This is considered by some as a reason to tax pension savings and other long-term investments more leniently. Diamond (2009) presents an illustrative example highlighting that a 30 percent tax on capital income only imposes a wedge of 3 percent between consumption today and consumption tomorrow (if the return is 10 percent) but that the tax wedge becomes 67 percent between consumption today and 40 years into the future. This should be compared with a 30 percent income tax, which implies a wedge of 30 percent between income today and consumption today.
Second, and perhaps most importantly, a fundamental restriction of the Atkinson-Stiglitz framework is that individuals are assumed to differ only along a single dimension. This implies that all inequality in capital income originates from individuals’ labor incomes (and labor earning abilities, in particular). This is perhaps the strongest reasons why the Atkinson-Stiglitz framework is not very useful to think about the optimal taxation of capital income. Later in this section, we discuss how heterogeneity in additional dimensions, for example, the form of inheritances received or differences in returns to investment, creates robust reasons to tax capital income.\(^{17}\)

Third, in their benchmark case, Atkinson and Stiglitz assumed that preferences are separable in consumption and leisure, implying (given that consumption in different time periods can be viewed as different commodities) that current and future consumption are equally substitutable for leisure. Hence there is no efficiency case for taxing future consumption via a capital income tax to offset the tendency for the labor income tax to induce substitution towards leisure.

### 3.2 Equity arguments in favor of taxing capital income

As we have already mentioned, the workhorse models of optimal taxation build on the assumption that individuals only differ with respect to their earnings abilities. This implies that differences in wealth across individuals solely derive from differences in skill and effort in the labor market. Some recent studies shed light on the fact that individuals differ in other important dimensions, which implies heterogeneity in wealth and capital income conditional on labor income. If these additional attributes correlate with individuals’ earnings abilities, taxes on capital become useful as indirect means to tax people with high ability.\(^{18}\) This relates to the fundamental information problem in the optimal income tax literature. If there is heterogeneity in capital income conditional on labor income, a tax on labor and capital income can achieve more redistribution than a labor income tax alone since capital income contains valuable information about individuals’ hidden earnings abilities.

#### 3.2.1 Heterogeneity in returns

In traditional models, individuals are assumed to earn the same risk-adjusted return on their investments. There is however a growing empirical literature documenting sizable differences in returns across individuals.\(^{19}\) If individuals with high labor earnings ability also have higher ability to generate a high return on their investment, because of access to social networks, information, or due to the economies of scale, another case where the Atkinson-Stiglitz results break down, is when investment affects the remuneration of low and high skilled labor differently. Pirttilä and Tuomala (2001) show that if an increase in investment leads to a decrease in the relative wage of low-income households, then a positive tax on capital income is desirable. The reason is that discouraging savings through capital income taxation reduces wage dispersion, which in turn makes progressive labor income taxation more efficient.

\(^{17}\) Another case where the Atkinson-Stiglitz results break down, is when investment affects the remuneration of low and high skilled labor differently. Pirttilä and Tuomala (2001) show that if an increase in investment leads to a decrease in the relative wage of low-income households, then a positive tax on capital income is desirable. The reason is that discouraging savings through capital income taxation reduces wage dispersion, which in turn makes progressive labor income taxation more efficient.

\(^{18}\) Banks and Diamond (2010), one of the background chapters to the Mirrlees Review, consider this to be one of the most compelling reasons to tax savings. Bastani and Waldenström (2019) document a robust relationship between wealth/capital income and skills in young adulthood using data from the Swedish military enlistment.

\(^{19}\) See Bach, Calvet and Sodini (2017).
forces working in favor of positive optimal capital taxation arise. For example, either highly skilled agents could achieve a higher rate of return by redirecting some of their time from labor supply into activities that raise their return on investment or they could simply be assumed to be inherently better investors. A small literature has also investigated the optimal tax implications of allowing for idiosyncratic investment uncertainty, a topic which we discuss in section 2.5.2 below.

3.2.2 Heterogeneity in preferences

In traditional models, individuals have the same time preferences for consumption. If high skill individuals have a higher taste for future consumption than low skill individuals, high skilled individuals will save more, even if they have the same income as a low skill individual. This implies that capital income should be taxed. The relationship between preference heterogeneity and optimal capital taxation has been analyzed in a few recent contributions (see Saez 2002, Diamond and Spinnewijn 2011 and Golosov et al. 2013). In this context, an important question is the following: should the government penalize individuals who prefer to consume tomorrow rather than today? Does it matter if the differences in savings behavior are the result of individual mistakes (for example, failure to estimate how much one values consumption at retirement)? These difficult questions have attracted some attention in the recent literature. For example, the issue of the treatment of pension savings in the tax system when individuals are subjective to self-control problems or cognitive biases, and the appropriate role of the government to deal with such issues has been analyzed by Moser and Silva (2017) and Hosseini and Shourideh (2017).

3.2.3 Heterogeneity in endowments (inheritance)

What are the implications of inheritance for the optimal taxation of capital? Cremer, Pestieau and Rochet (2003) analyze a model where individuals have the same preference for saving, but instead differ in terms of their endowments/inheritance (assumed to be exogenous). If there is a positive correlation between endowments and earnings ability (for instance, due to a genetic correlation in earnings ability or an inter-generational transmission of economic opportunities) this implies that two individuals with the same labor income, but with different abilities, also differ in terms of the amount they can consume because of their endowment. This implies that these two individuals have different demand functions for goods (including future consumption) and provides a motivation for taxing capital income. This argument relies on the government not being able to observe inheritance; otherwise, all differences in initial endowments could be eliminated through confiscatory taxation.

---

20 See Gahvari and Micheletto (2016), Kristjánsson (2016), and Jacobs et al. (2018) for recent contributions.
21 The discussion here complements earlier surveys on the topic, such as Cremer and Pestieau (2011).
22 See also Brunner and Pech (2012) for an extension.
Farhi and Werning (2010) allow, in contrast to Cremer et al., bequests to be taxed analyze the properties of an optimal estate tax. They focus on two generations where all capital (and inheritance) derives from the work efforts of the first generation. Individuals in the second generation will be heterogeneous in terms of the bequests they receive depending on who their parents are. In such a framework, they show that if one considers the welfare of the parents (those who give bequests) but not the welfare of the children (those who receive bequests) then the inheritance tax should be zero, essentially in line with the AS theorem. When also considering the welfare of those who receive bequests, then they find that it is optimal with a progressive (negative) inheritance tax which subsidizes inheritances, but with a degree of subsidization that decreases in the size of the inheritance. If it is not possible to subsidize inheritance for some exogenous reason, the degree of subsidization will be zero for all but the largest inheritances, which should be taxed. The usefulness of the progressive estate tax is that it equalizes the bequests that people receive, which raises the welfare of the second generation.

Farhi and Werning (2013a) build upon Farhi and Werning (2010) and highlight the fact that parents differ in terms of how altruistic they are towards their children. This creates inequality between children depending on the preferences of their parents to bequeath to their children, even conditional on the earnings abilities of the parents. The optimal estate tax considers that inheritance taxation discourages labor supply activity of parents, while it levels the playing field of the child generation. In comparison to their earlier study, Farhi and Werning (2013a) find it can be optimal to tax inheritance if the principle of equality of opportunity carries sufficient weight in the objective function of the tax designer, which they argue could be an explanation why inheritance taxation exists in many countries.

A restrictive assumption in the analysis of Farhi and Werning (2010, 2013a) is that they examine a two-period model, with one generation of parents who give bequests and one generation of children who only consume. Piketty and Saez (2013) study a more realistic setup where each generation both gives and receives bequests. This implies that those who bequeath to a greater extent are those who have inherited in the past. In addition, their analysis considers a correlation in earnings abilities across generations, which implies that those who receive large inheritances are more likely to also be individuals with a high earnings ability. Similar to Farhi and Werning (2013a), the results of Piketty and Saez (2013) derive from the assumption that individuals differ not only in terms of their earnings abilities but also in terms of their altruism towards their offspring. The fact that earnings abilities are no longer the unique determinant of life-time resources is the main reason for the finding of positive inheritance taxation in their model.\(^{23}\) \(^{24}\)

---

\(^{23}\) For mathematical reasons, Piketty and Saez restrict attention to a linear inheritance tax. Their analysis is therefore not informative about optimal progressive inheritance taxation, analyzed by Farhi and Werning.

\(^{24}\) De Nardi and Yang (2016) quantitatively analyze inheritance taxation in the US. In their model, individuals are born with different circumstances, both with respect to inheritance and in how much their parents have invested in their human capital (alternatively, allowing for a genetic correlation in ability across generations). They find
To sum up the insights from the above-mentioned papers, we can conclude that a positive correlation between earnings ability and inheritances received as well as the circumstantial nature of inheritance due to parental altruism heterogeneity are arguments in favor of inheritance taxation. However, the “double-dividend” of giving (referring to the fact that inheritance provides utility not only to the donor but also to the recipient) represents a serious challenge in assessing the welfare effects of inheritance taxes that has not yet been resolved. The benefit to the recipient is an externality, and therefore motivates to tax bequests less than other consumption goods.\textsuperscript{25} In conclusion, we also would like to mention that heterogeneity in endowments is also a motivation for wealth taxation, in particular if there are information constraints or political constraints that prevent the optimal use of inheritance taxation.

3.3 Efficiency arguments in favor of taxing capital income

3.3.1 Reducing distortions on human capital accumulation

One of the most important objections to the Chamley-Judd analysis concerns its assumption that only capital accumulates over time. In economies with progressive income taxation, an equally serious concern should be to provide incentives for individuals to invest in education, exert effort on the job, and advance in their careers. That is, the accumulation of human capital can be just as important as the accumulation of physical capital. Jacobs and Bovenberg (2010) analyze the role of human capital accumulation for the desirability of taxing capital income. They find that, in a model where not all education investments are verifiable by the government, and therefore cannot be directly subsidized, a positive tax on capital income serves to alleviate the distortions of the labor tax on human capital accumulation. The intuition is that labor taxes encourage individuals to substitute human by financial assets. The optimal capital tax trades off the distortions on the overall level of total savings and the composition of saving (human vs financial assets), which can lead to positive capital taxation. Importantly, the result in Jacobs and Bovenberg does not depend on non-separable utility, heterogeneous preferences, or financial market failures. Since that study, there has been a surge of papers emphasizing the importance of considering human capital accumulation in optimal tax analysis. Stantcheva (2017) is a recent contribution that further discusses this strand of the literature.

3.3.2 Taxing economic rents

The literature on optimal income taxation has almost exclusively analyzed how the so-called “normal” return to savings should be taxed (such as the return to an average investment or the yield of a government bond). As acknowledged above, returns are likely to be heterogeneous across individuals. If an individual earns a return on an investment that is greater than the normal rate of return, this is an \textsuperscript{excess}

\textsuperscript{25} See Boadway and Cuff (2015) for a recent discussion.

that in the long-run equilibrium, estate taxes of inheritances over a certain threshold have small or insignificant effects on the capital accumulation of the economy, but can deliver large welfare gains for newborns who do not know in which economic environment they will grow up, while generating large welfare losses for the very rich.
return”. If excess returns reflect chance events (and not factors over which individuals exert control), they are viewed as “economic rents”, which can be taxed without distortions. On the other hand, if excess gains are the result of productive economic activity, taxing excess returns entail distortions.

An important difference between a tax on the stock of capital (a wealth tax), and a capital income tax is the taxation of excess returns. If the normal return on an investment is 5 percent, a capital income tax of 20 percent is comparable to a wealth tax of 1 percent. However, for individuals who receive a return greater than 5 percent, they will have to pay tax on the excess return under capital income taxation, but not under a wealth tax.

Most of the theoretical research on the optimal design of capital taxation starts from models where a single asset is used as vehicle for consumption smoothing and to fuel investments in the economy. To usefully analyze the taxation of excess returns, one can consider a framework where individual investors choose between risky and safe assets. The standard model to analyze this issue goes back to Domar and Musgrave (1944) which was further generalized by Stiglitz (1969). The main message is that, under certain conditions, investors who are faced with a capital income tax with full loss offset (implying that losses can be deducted from the tax base) can change their choice between risky and safe assets in a way that generates the same combination of risk and return as in a no-tax situation. The reason is that the capital income tax provides insurance that allows agents to increase the risk of their portfolios, mitigating the effect of the tax on the after-tax return.

It is a difficult, but important, empirical exercise to determine to which extent the taxation of excess returns represents taxation of economic rents and to which extent it represents distortionary punishment of highly skilled investors. Most economists would probably agree that it is desirable to tax excess returns, and the academic discussion has mostly centered on the whether or not to tax the normal return to savings. The Mirrlees Review (Mirrlees et al. 2011) recommends taxing excess returns but not taxing the normal return to savings. Notably, this policy advice goes against their background report, Banks and Diamond (2010). The main argument in Mirrlees et al. (2011) is that the taxation of the normal rate of return violates principles of neutrality in the tax system. However, the purpose of the tax system is not to achieve neutrality, but to maximize social welfare. Thus, capital income taxation must be judged by how it interacts with the desire to redistribute income at the lowest efficiency cost. Few studies, however, have analyzed the optimal taxation of capital income in models featuring idiosyncratic investment risk and ex-ante heterogeneous individuals. An exception is Spiritus and Boadway (2018). These authors build upon the works of Gordon (1985), Christiansen (1993) and Schindler (2008), and develop

26 See also Atkinson and Stiglitz (1980).
27 Norway allows, since 2006, a tax-free normal rate of return on investments in stocks. According to Sørensen (2005) this system does not distort firm’s marginal investment decisions and how these investments are financed within the firm. This conclusion has been criticized by Lindhe and Södersten (2012) who suggest that neutrality of this kind is not fulfilled when returns to investments largely are determined by international capital markets.
a general framework to analyze the question whether the exemption of the risk-free component of capital income (normal rate of return), as recommended by the Mirrlees Review, continues to hold when some assets earn excess returns. In the standard mean-variance portfolio framework where returns to private investment satisfy constant returns to scale (and hence, do not include rents), taxing excess returns and exempting normal returns (as in the Mirrlees review) is optimal. If, on the other hand, there are returns to scale in private investments, the normal return to capital should be taxed when returns to scale in investment are decreasing, and subsidized when they are increasing. Notice that these results apply in a model where all other arguments for positive capital taxation discussed in this section explicitly have been ruled out.

3.3.3 Correcting capital market failures

Atkinson and Sandmo (1980) is a seminal study of capital taxation in an OLG framework where each generation lives for two periods, working in the first, and being retired in the second. In this setting, they found that it can be desirable to tax capital income for a reason related to the well-known property of OLG models, namely that the economy does not always reach its full production capacity since current generations do not consider the effects of their savings on future generations (each generation lives for a finite period whereas the economy lives forever). This dynamic inefficiency can effortlessly be corrected if the government is free to issue public debt or is allowed to use age-dependent lump-sum transfers. However, when there are restrictions on the use of such instruments, a positive capital tax can be desirable as it enables redistribution between different generations.

Atkinson and Sandmo demonstrated that a positive tax on capital income can be desirable in order to induce agents to save more if the income effect on savings is sufficiently strong. Moreover, a positive capital tax can finance tax reductions on labor, which can be a way to make younger generations save more. At the end of the day, it is however unclear how large of a role intergenerational redistribution issues should play when designing taxes on labor and capital. There are other ways to redistribute between generations that are more effective, for example by adjusting the pension system. The Atkinson-Sandmo framework also only considered a model with a representative agent. Later studies have analyzed OLG models with redistribution motives both between and within generations (due to skill heterogeneity), which makes the policy implications of dynamic inefficiency less clear.28

3.3.4 Correcting insurance market failures

A well-known situation, in which the models of Atkinson-Stiglitz and Chamley-Judd lead to a positive capital income tax, is when future earnings are uncertain. In a perfect market, individuals would be able to handle the prospect of an uncertain income by borrowing in periods with low income and pay back

these loans when incomes have recovered. The problem is that the market is not perfect, individuals cannot always borrow, and there are many risks that are difficult or impossible for individuals to insure themselves against. This can give rise to precautionary savings, where individuals save in periods with high income to secure their consumption in periods with (unexpected) low incomes.

Aiyagari (1995) considered an infinite-horizon model where individuals are borrowing-constrained, face uninsurable uncertainty about their future wages, and only decide about how much to consume in each time period. The borrowing constraints prevent individuals from insuring against idiosyncratic shocks, implying precautionary savings. In this setting, Aiyagari showed that the incompleteness of markets could lead to an over-accumulation of capital, which motivates positive capital income taxation to reduce this over-accumulation and move the capital stock closer to the level prevailing with complete markets.29 This is related to the capital market failure that happens in OLG models, discussed above. However, in the Aiyagari economy, the government cannot use public debt to correct the market failure and move the economy to the optimal level of capital accumulation. That there could be arrogate welfare gains from reducing the capital stock, seems, however, to be somewhat implausible.

The distinction whether or not individuals face deterministic or stochastic productivity profiles over their life cycle also matters for the desirability to tax savings in models analyzing nonlinear income tax systems. If individuals face uncertainty regarding their future productivity, individuals might self-insure through their savings. This motive to save implies a negative impact on labor supply. The reason is that individuals tend to save “too much and will bring the same amount of savings into the future, irrespectively of if they realize a high or a low productivity in the future, which has a negative effect on labor supply in both states. The provision of insurance over the life cycle in response to uncertain productivity is the focus of the so-called New Dynamic Public Finance literature (see Golosov et al. 2006).30 Uncertainty about labor productivity seems, however, to be of secondary importance to the taxation of capital income, as suggested by Farhi and Werning (2013b) and Bastani, Blomquist and Micheletto (2013).31

### 3.3.5 Reduce distortions on labor supply

In early life cycle models of optimal capital taxation, it was common to analyze models where individuals work only in the first period of life. The subsequent literature has analyzed the optimal taxation of capital income in lifecycle models where workers work in multiple periods. In such a setting, age-

---

29 Borrowing constraints are common components of modern models used to analyze capital income taxation, such as Conesa et al. (2009).

30 Two of the most important papers in this literature are Albanesi and Sleet (2006) and Golosov et al. (2016).

31 Sometimes scholars refer to the result that capital taxes should be “zero in expectation”. This result, shown by Kocherlakota (2005), refers to the fact that in these models, capital taxation over the total population in each time period balance each other out, but there will be positive capital taxes on those who have an incentive to over-accumulate capital. An interesting observation in this context is that it is typically the poor who find insurance valuable and tend to “save too much” and therefore should have their savings discouraged by the tax system. In this context, capital taxes are only used to deter shirking, and not to perform any actual redistribution.
dependent labor income taxes become desirable due to age-specific labor supply behavior. Erosa and Gervais (2002) show that if age dependent labor income taxes are not available, and (realistically) individuals’ life-cycle productivity profiles are not flat, positive capital income taxes are desirable because they can serve as a substitute for age dependent labor income taxation. The intuition is that if consumption is a stronger complement to leisure later in life, as compared to earlier in life, it is optimal to tax savings in order to boost labor supply and reduce the distortions associated with labor income taxation. This is reminiscent of the classic result by Corlett and Hague (1953) recommending that goods complementary to leisure should be taxed. However, the reason for a positive tax on capital income in life-cycle models survives even if the utility function is weakly separable between consumption and leisure, making it an inherently dynamic result. The life-cycle elements in labor supply is one of the essential features of the economy analyzed by Conesa, Kitao and Krueger (2009), who find a positive and sizable optimal tax on capital income in their simulations calibrated to fit the US economy.

The early literature analyzing capital income taxation in dynamic frameworks considered economies with a representative individual and with a focus on linear (proportional) tax instruments (and in the case of Erosa and Gervais 2002, a representative individual within each generation). The following literature has analyzed richer dynamic models where agents are heterogeneous in skills, work in multiple periods, and face deterministic or stochastic productivity profiles over their life cycles. The goal of the social planner in these settings is to achieve redistribution or insurance at the lowest efficiency cost. These models imply that, in general, it is optimal to tax savings since individuals tend to react to progressive income taxation by working less and consuming their savings. This result is similar to the motivation for taxing savings that occur in life-cycles models of the type studied by Erosa and Gervais (2002). Here, consumption late in the life cycle is more complementary to leisure than is consumption early in the life cycle. However, a key difference is that the taxation of savings in dynamic Mirrlees models arises from the desire to redistribute income (or provide insurance) through nonlinear income taxation and the taxation of savings enables to counteract the distortions associated with income taxation and thereby perform redistribution at a lower efficiency cost. To see this most clearly, consider the case of a high wage thirty-year-old. If this person anticipates having a high wage also when in his/her fifties, he/she might choose to work less when in his/her fifties. The benefit of doing so would be that, when this person is in his/her fifties and is working less, he/she would have the same income as a low-wage person working full time and, if there is progressive income taxation, qualify for a lower tax burden. However, the high-wage person would save a larger amount as compared to the low wage person, and

---

32 These effects are present in any life-cycle model, not only models where generations overlap (OLG).
33 See propositions 3.2 and 3.3 in Erosa and Gervais (2002).
34 There are some exceptions, such as the paper by Ordover and Phelps (1979), that considered the optimal nonlinear taxation of labor and capital income in an OLG model where agents are heterogeneous in skills, as in Mirrlees (1971). However, in these papers, individuals typically supplied labor in the first period and were retired in the second.
therefore be able to consume more. Taxing savings implies that such reduction in labor supply in response to progressive income taxation becomes less attractive.

The desirability of taxing savings to improve the efficiency of the tax system crucially depends on the sophistication of the income tax available to the government. If the government could impose different taxes on individuals in different ages then high wage individuals in their fifties could be provided with age-specific incentives to supply high amounts of labor without the need to disrupt the incentives of thirty-year olds. However, in contrast to the analysis of optimal capital taxation in representative agent models, the presence of within-generation heterogeneity makes it desirable to tax capital income even if the labor income tax is allowed to be age dependent. If the labor income tax is even more sophisticated, however, so that it can be both age and history dependent (depending on the present and past labor incomes of an individual) then the gains of taxing savings to combat labor income tax distortions becomes smaller or disappear completely. In the above example, an individual supplying a high income when young could be rewarded if he/she continues to earn a high income as middle-aged if the income tax is history dependent, mitigating adverse effects of savings on future labor supply.

4. Specific capital taxes: Theoretical and empirical considerations

As discussed above, most of the theoretical research on the optimal design of capital taxation emanates from models where a single asset is used as vehicle for consumption smoothing and to fuel investments in the economy. In practice, individuals invest in a variety of assets that have special properties, such as financial assets, pensions, housing and closely-held businesses. In this context, an important aspect that is often overlooked is the distinction between wealth and capital. Not all savings and wealth is productive and facilitates physical investment.

The conventional wisdom among economists is that, when there are multiple assets categories, the returns to investment in these assets should be taxed uniformly. This is sometimes referred to as ‘neutral’ capital income taxation. This avoids not only distortions in consumer choices (to the extent that assets provide consumption benefits), but also avoids distortions in production. The rationale for uniform capital income taxation to avoid production distortions is based on the so-called production efficiency theorem developed in the seminal contribution of Diamond and Mirrlees (1971). In short, the result relies on the observation that taxing different input factors in different ways distorts production, and these distortions, in the end, manifest themselves in the form of different consumer prices. As the effect of the input tax differentiation can be replicated by using differential taxation of final consumption goods, eliminating the differential taxation of production inputs and replacing it with differential taxation of final commodities can generate a Pareto-improvement, as this will increase total output produced in the

---

35 In fact, with deterministic productivity profiles, zero taxation of capital income is optimal if the income tax is history dependent.
economy. The Diamond and Mirrlees analysis relied, however, on some rather specific assumptions, such as that pure profits can be fully taxed. If this is not the case, it is desirable to impose higher tax rates on input factors used in sectors characterized by imperfect competition or assets where prices are above long-run marginal costs that reflect economic rents (Dasgupta and Stiglitz 1972). It should also be pointed out that the production efficiency theorem should not, however, be interpreted as an argument against taxing capital income in general, as clarified by Diamond and Saez (2011).

The most important argument in favor of uniformity, however, is that it is an important benchmark case and usually regarded as a useful principle in modern tax systems. It is very difficult to know what an optimal capital income tax differentiation would look like since the optimal differentiation would depend in a complex way on a host of elasticities of substitution in consumption and production and properties of the financial markets. Moreover, uniform taxation is also valuable to suppress attempts of special interest groups to pressure politicians to modify the tax system.

4.1 Wealth taxation

A wealth tax is a tax on the stock of capital. In principle, wealth taxes would not be needed if all income sources that form the basis of wealth would be taxed. Since inheritance is the intergenerational transmission of wealth, the same argument could be made about inheritance taxation. Nonetheless, many countries do tax both wealth and inheritance. One main motivation for this is that the wealth distribution is a function of not only the present design of the tax system, but also the design of the tax system in the past, as well as present and past opportunities for tax evasion, tax arbitrage and tax planning. In a sense, wealth and inheritance taxation can be viewed as ways to compensate for the inability to tax income optimally in the past. This can be due to an inability of the government to tax certain kinds of hard-to-observe income, but also due to political or administrative failures.

The wealth tax base is not easy to define. According to the official wealth definition in the UN’s System of National Accounts, private wealth includes not only real estate, bank deposits, bonds, corporate equity etc., but also all funded insurance savings in life insurance and occupational pension schemes. Including the latter in the tax base not only creates administrative challenges but can also imply an unreasonable tax burden for many low-income households. However, excluding insurance and occupational pension schemes from the tax base distorts investment choices. One major problem is the valuation of assets, especially the equity of non-listed firms. In the absence of secondary market prices, these firms have to be valued based on accounting information and errors in the valuation creates both uncertainty and potential liquidity problems. In an attempt to respond to such problems, most countries introduced reliefs, and even total exemptions, on business assets. While these measures alleviated some problems, they represented a departure from the conceptually advantageous broad-based feature of the

36 Sørensen (2007) provides a more elaborate discussion of capital tax neutrality and open-economy issues.
general wealth tax. Finally, international mobility of capital is seen as a severe problem for the wealth tax. Even if we still do not know much about how important this constraint is in reality, cross-border capital flight represents a credible criticism against wealth taxation (and capital income taxation). Nonetheless, wealth taxation might be one of the few ways to tax business assets representing the bulk of the wealth of the very wealthiest in society.

In ongoing work, Guvenen et al. (2017) analyze wealth taxes in a model where individuals have different returns to their investment. They propose that there could be welfare gains associated with shifting from capital income taxation to wealth taxation. The argument is that when only capital income is being taxed, the burden of taxation falls disproportionately on high skilled investors, whereas passive and less successful investors avoid taxation. An interesting conflict therefore arises between redistribution (those who generate high returns have a high earnings ability) and efficiency (taxing capital income can lead to reduced investment among high skill investors). This presumes that excess returns are created by productive activities and not luck or circumstance.

There is not much empirical research on the efficiency cost of wealth taxes. The main reason is the lack of adequate data and credible identification strategies. Housing wealth constitutes the bulk of most household portfolios, and is considered to be fairly insensitive to wealth taxation (apart from capitalization effects). Entrepreneurial activity and business wealth are perhaps what economists are mostly interested in but is imperfectly covered in most wealth databases and sometimes not even part of the tax base.

Most empirical studies examining behavioral responses to the wealth tax analyze taxable wealth rather than the economically more relevant total marketable, market-valued wealth. In other words, these studies capture how owners adjust their investments according to how they are taxed (reporting effects) rather than the allocation of real physical investment (see Brühlhart et al. 2017 and Seim 2017 for two recent contributions). Jakobsen et al. (2018) is the perhaps most ambitious attempt so far to identify the real effects of wealth taxation by analyzing behavioral responses of taxpayers. Using a rich Danish administrative register dataset and the abolition of the Danish wealth tax in 1997, they find that the behavioral effects of the wealth tax on wealth accumulation were small in general, but large among very wealthy households. The order of magnitude of the estimated effects indicate fairly notable efficiency costs of the wealth tax, although the estimates should be interpreted with some caution due to the uncertainties in the historical data series.

## 4.2 Taxes on personal capital income

Personal capital income refers to the return on a person’s capital stock, which includes interest income, dividends, realized and unrealized capital gains. In similarity to wealth taxation, the taxation of capital gains represents a serious challenge for the tax system. In contrast to the taxation of capital income in
the form of dividends or interest income, the problem with the taxation of capital gains is that they should ideally be taxed when they occur, but are in most countries taxed only when they are realized in a market transaction. This gives rise to a lock-in effect where individuals hold on to their assets in order to not trigger a tax payment. This lock-in effect can be mitigated by taxing assets continuously by computing an imputed rate of return on the asset. The downside is that this is a tax payment that does not correspond to an income stream, and hence generates liquidity problems. One solution is discussed by Auerbach (1991) who proposed to tax capital gains upon realization, but charging an interest on past gains when realization finally occurs, effectively eliminating the incentives to postpone realization. In the case of capital gains pertaining to housing, an annual property tax based on assessed market values can be used to continuously tax latent capital gains, thereby reducing the need to tax everything upon realization and avoiding some of the lock-in effect. We discuss property taxation in more detail below. It is fair to say, however, that the lock-in effect is hard to avoid completely, and is often used as a pragmatic argument to keep the capital income tax rate relatively low.

A major argument for taxing capital income less than labor income is that capital income is more responsive than labor income. Kleven and Schultz (2014) find that the elasticity of capital income is two to three times as high as the elasticity of labor income using Danish data. Rydqvist, Spizman and Strebaev (2014) show how postwar dividend tax policies have been pivotal to explain the decline in individual direct stock ownership while stock holdings of financial institutions have surged, especially within the context of pension plans.37

For both equity and efficiency reasons, the optimal taxes on labor and capital income are likely to be a fully nonlinear function of both labor and capital income. In other words, individuals with low and high labor income should face different capital income taxes and optimal capital income taxes are likely to be progressive, namely, the capital income tax rate is different for individuals with low and high levels of capital income. While there are some important exceptions, such as asset-testing in the context of welfare programs, and progressive estate taxation, actual tax systems tend to make limited use of such inter-dependencies and non-linearities. One reason is the problems of tax arbitrage. If one tried to tax savings through a nonlinear function, there would be large incentives for someone with a high marginal tax on savings to ask a friend or a relative with a lower marginal tax on savings to save for him. This is essentially the same argument that prevents the nonlinear taxation of commodities, namely, the difficulties for the government to observe and verify personal consumption levels.

37 There is a large literature on how capital taxes matter for the composition of household portfolios. For example, Agell and Edin (1990) showed how increases in the tax on bank interest income led to an increase in the share of stock investments among Swedish households. See Poterba 2002 and Campbell 2006 for reviews of the earlier literature.
The US and many other countries adopt some form of the so-called **comprehensive income tax** where
the sum of labor and capital income is taxed together according to a nonlinear tax schedule.\(^{38}\) A benefit
of the comprehensive income tax is that it taxes all sources of income, at the margin, at the same rate,
which reduces incentives for tax planning (cross-base shifting). However, according to optimal tax prin-
ciples, taxing labor and capital income at the same rate is sub-optimal.

A more flexible system is the Nordic so-called **Dual Income Tax (DIT)**, which combines the progressive
taxation of labor income with the (separate) proportional taxation of capital income.\(^{39}\) From an optimal
tax perspective, the DIT has the desirable feature that the capital income tax rate and the labor income tax rate
can be made different for high-income earners. At the same time, an optimal dual income tax must consider the possibility for individuals to shift between the labor and capital income tax bases.\(^{40}\) The latter is usually presented as an argument in favor of not making the difference between the top marginal labor income tax rate and the proportional capital income tax rate too large. In particular, it is an argument for a positive tax on capital income.\(^{41}\)

Variants of the DIT have been used in the four Nordic countries for about 30 years.\(^{42}\) It should be recognized that several of the arguments for the DIT are pragmatic in nature. Sørensen (1994) describes the historical background for the introduction of dual income taxation in the Nordic countries and discusses a number of practical and administrative arguments for the DIT. The most important arguments are the fact that the capital income tax is levied on nominal rather than real income (implying that the effective tax rate is higher than the statutory rate when the inflation rate is positive), and the point that a relatively low capital income tax rate makes it easier to subject all forms of capital income (including capital gains) to a uniform tax. The idea with this “uniformity” or “neutrality” of the capital income tax was that a uniform proportional tax rate would apply to all asset types and holding periods, allowing

\(^{38}\) Such systems are based on the notion that it is the sum of all incomes that is relevant to the well-being of individuals. In addition, having individuals with the same total income pay the same income tax can be argued to respect the principle of horizontal equity. In practice, however, the comprehensive income measure that is available to tax authorities is seldom a complete account of all the sources of income that are relevant to an individuals’ welfare, as there are sources of income that are not observable, such as the intra-family transfers and unrealized capital gains.

\(^{39}\) The progressive taxation of labor income is administratively feasible by virtue of the now widespread use of third-party reporting of income to the tax authority (Kleven et al. 2011).

\(^{40}\) Alstadsæter and Jacob (2016) have found evidence of income shifting in Sweden among business owners by using a reform in 2006 in which both the tax differential and the amount eligible for reclassification increased. Pirtilä and Selin (2011) study similar questions in Finland and find evidence of income shifting, especially among the self-employed.

\(^{41}\) See Christiansen and Tuomala (2008) for a theoretical argument in favor of taxing capital income due to the possibilities for income shifting. See also Selin and Simula (2017) for a recent analysis of the social welfare effects of income shifting.

\(^{42}\) Denmark was the first to introduce the dual tax system in 1987. Thereafter, Sweden did it in 1991, Norway in 1992 and, finally, Finland in 1993.
deductions for capital losses and capital expenses, thereby minimizing incentives for tax planning, tax arbitrage and other distortionary activities.

However, the uniformity turned out to be difficult to uphold due to political pressures from special interest groups to implement tax changes catered to specific groups in society. In Sweden, for example, differentiated tax rates and changes in the method used to calculate taxable returns implied that the effective tax rate could differ across types of capital incomes.\textsuperscript{43} Policymakers have motivated the departure from uniformity in different ways, mainly by pointing to the need for special exemptions to promote business activity among small and middle-sized firms.

Appropriately calibrating the marginal tax rates on labor and capital income in the dual income tax system is an important task. As discussed above, optimal marginal tax rates on labor and capital income are not necessarily the same. The appropriate calibration of these tax rates depends on both equity and efficiency factors. The efficiency considerations are governed by the elasticities of the labor and capital income tax bases as well as the elasticities of cross-base income shifting. From an equity point of view, the calibration of optimal tax rates depends on the correlation between the labor and capital income tax bases and the welfare of individuals. Saez and Stantcheva (2018) make an attempt to calibrate these optimal tax rates. In the case where they allow for nonlinear taxes on labor and capital income, and the elasticity of capital income is higher than the elasticity of labor income, the structure of capital income tax rates appear relatively flat, and are lower than labor income tax rates. This suggests that the dual income tax system, with a proportional tax on capital income, at a lower rate than the marginal tax rate on labor income, might not be that very far from optimal. However, further research is needed to quantify the welfare losses associated with adopting simple tax systems, such as the dual income tax system, using realistic measures of cross-base income shifting.

It is important to recognize that the elasticity of the capital tax base is partially determined by the opportunities individuals have to conceal their wealth abroad, which can directly be affected by policy. Some research studies have found evidence of capital flowing back from tax havens as a result of the information treaties, and particularly those that are signed at the multilateral level. Johannesen and Zucman (2014) study the effect of bilateral treaties regarding the reporting of banking transactions. Their main finding is that tax evaders seem sensitive to the risk of exposure, but instead of repatriating, shifted their funds to tax havens that were not covered by the treaties. Slemrod et al. (2017) analyze how a series of US initiatives have affected tax evasion of US citizens presumably hiding assets around the

\textsuperscript{43} Specifically, dividends and capital gains are taxed at different rates depending on whether the companies are listed on the stock exchange or not. Some investments in special “investment accounts” are taxed based their imputed, rather than actual, return, making them more like a tax on the stock of financial wealth.
world. Although results are preliminary, individuals reporting foreign assets increased by 20 percent, but the largest effect came from voluntary compliance outside the control initiatives.

4.3 Property taxation

A specific type of investment is land and buildings. “Property taxation” or “real estate taxation” is an annual tax on real property where the tax base may be land or buildings, or some combination of the two. The tax is usually based on an assessment of the market value of the property.

An investment in property is similar to an investment in a business asset. In the case of housing, the owner receives a return either in the form of an income stream obtained by renting out the house or a consumption stream obtained by using it as a personal accommodation. In the former case, the income accruing to the proprietor is observable, and can be directly taxed. In the latter case, the tax authority needs to make an estimate of the value of the consumption the investment generates to its owner, which is referred to as the “imputed rent”. A common feature in many countries is that housing is subsidized in many ways. Poterba (1992) discusses these issues in the US and calculates the net-of-tax income from owning a house, considering it a business asset generating real economic profits, and discusses how these profits should be taxed in order to avoid efficiency costs associated with the subsidization of housing (see also Poterba and Sinai 2008). Gervais (2002) quantifies the welfare losses associated with the preferential tax treatment of housing in the US using a calibrated life-cycle model. He finds that individuals at all income levels would gain by the proper taxation of imputed rents and a restriction of the deductibility of mortgage interest.

The theory of optimal nonlinear income and commodity taxation can be used to assess the role of the taxation of housing in an optimal tax system. According to the Atkinson-Stiglitz theorem, the imputed rent should be taxed in the same way as all other consumption goods, unless there is a good reason not to.44 One such good reason could be if high skill individuals reduce their labor supply in order to perform home improvements that raise the value of the house. In this case, housing taxation becomes an indirect way of taxing leisure, increasing the attractiveness of work, which could mitigate the distortionary costs associated with progressive income taxation.45 Another reason why one would like to deviate from taxing properties in accordance with other goods would of course be if there are externalities. Some have argued that there are positive externalities if people take good care of their houses, as

44 The analysis of AS is not directly applicable to housing taxation as it neglects the durable aspect of housing consumption, which suggests that a dynamic framework is needed to analyze it. Köhne (2018) develops a dynamic framework of optimal commodity taxation where some goods are durable. In an application to housing, he finds that housing investment should be taxed at a higher rate than nondurable consumption.

45 However, a more efficient tool to mitigate these distortions could be to provide tax breaks for home improvement services purchased in the market, See Köhne and Sachs (2019).
it provides a benefit to other people, and may result in better neighborhoods. Others argue that marginal quality improvements in housing produce negative externalities if individuals compare their housing consumption with others (that is, status-effects or envy, see Alpizar, Carlsson and Johansson-Stenman 2005, and Aronsson and Mannberg 2015).46

In the context of property taxation, it is worth noticing that if the value of a house mainly reflects the value of the land upon which it was built, and the land value reflects economic rents, then additional taxation of housing beyond that to achieve uniformity with respect to other goods is warranted, or even better, a separate tax on the (unimproved) value of land.47 Finally, it is also important to distinguish between properties such as housing, and commercial properties that are used as inputs in production. To avoid distorting production choices, commercial properties should be taxed in the same way as other inputs in production.

The correlation between individuals’ capital ownership and their ability to generate income is one motivation for property taxation. The strong concentration of real estate capital in the upper part of the income distribution could be a particular motivation for a progressive property tax. Another factor that could motivate a more progressive tax on real estate would be if wealth directly enters individuals’ utility functions, as real estate wealth constitutes the bulk of total wealth for most households.48

In the public debate, the property tax is a recurrent theme. Economists embrace it because it is efficient, while the public usually are less positive. In the US, so-called “property tax revolts” have erupted recurrently since the 1970s, often associated with middle-class homeowners protesting against the tax and many times successfully convincing policymakers to amend policies (Martin 2008).

Salience is another potential determinant of the support for property taxation.49 Individual homeowners themselves often have the responsibility to pay the tax to the tax authority, in contrast to other taxes, such as income taxes, which often are withheld at source, and therefore less visible to taxpayers. Cabral

---

46 A large literature has found that consumption goods are not only valued based on their absolute qualities, but to a large extent how they compare to the consumption of others. Alpizar et al. (2005) found that housing is consumption good where such ‘relative consumption concerns’ are the strongest. For instance, it is likely that a person could achieve a higher utility living in an expensive house in an area where the average price of housing is low, as compared to living in an equally expensive house in an area where the average price of housing is high (ceteris paribus).

47 Land is immobile and is the canonical example of a tax base for which price appreciation is independent of personal effort and almost exclusively determined by demand and supply. For example, a new public transport facility that reduces commuting time in a certain area will result in higher land prices in that area. Still, this capitalization effect also goes for the buildings on the land, and therefore one usually tax not only the land, but the entire real estate.

48 Progressive property taxes exist in several countries, for example, Denmark, Finland, France, Germany and Norway. Progressivity can appear in different forms. There can be a progressive tax schedule, or a basic deduction in combination with a proportional tax rate.

49 For contributions emphasizing the importance of tax salience, see Chetty, Looney and Kroft (2009) and Finkelstein (2009).
and Hoxby (2012) study the relationship between salience and the level of property taxation in the US by comparing US states where the degree of salience varies as a function of technical features of the tax collection. They find that the salience of the property tax could be one important factor explaining why it is so unpopular.

Liquidity problems are another possible explanation for the low popularity of property taxation, but one that has received less attention in the academic literature. There are practical tools that can deal with such issues. A limitation rule can mitigate the tax burden for people with low income (introducing, however, adverse labor supply incentives in certain income ranges). A dampening rule can be used to smooth tax payments in periods with soaring real estate prices. Another possibility would be to allow postponement of (some of) the property tax bill until the time the property is sold, but adding a market interest rate to the postponed tax liability to maintain the present value of the tax payment, thereby avoiding lock-in effects.

4.4 Inheritance taxation

In policy discussions about inheritance taxation, a few specific efficiency considerations frequently present themselves. First, inheritance can have negative effects on government revenue if those who receive an inheritance work less (an income effect). From this perspective, inheritance taxation can provide additional positive effects on government revenue beyond the direct mechanical effect (this is sometimes labelled a positive “fiscal externality”). Second, taxing inheritance may make it less attractive for parents to work if a motivation for working is the possibility to transfer resources to the next generation. Third, to the extent that bequests are accidental, taxing them is efficient. The literature on optimal inheritance taxation is not conclusive, and further work is needed to understand how these aspects affect optimal tax analysis. Moreover, further research is needed in order to understand why individuals bequeath their wealth. This is relevant both when assessing welfare effects and when assessing the effects on labor supply. Finally, we may note that there are other transfers to children that are not taxed, such as human capital investment. This means that inheritance taxation might distort parent’s decisions about how to invest in their children.

A recurrent issue with the inheritance and gift tax is how business assets should be treated, in particular those relating to the generational succession of family firms. Many countries have introduced reliefs for these asset types. A common motivation for such reliefs is the liquidity problems that may arise which potentially may require heirs to sell their shares to finance the payment of the tax (and then potentially

50 See Kindermann et al. (2017) for a recent contribution assessing the importance of this effect.

51 However, Blumkin and Sadka (2004) and Cremer, Gahvari and Pestieau (2012) question the desirability of 100 percent taxation of accidental bequests.
incur additional taxes when latent capital gains become realized). Valuation problems are another common motivation for implementing special reliefs for inherited business assets. 52

To empirically analyze the economic consequences of inheritance taxation is difficult and there are few studies available. One of the major challenges is to distinguish actual capital accumulation effects from tax planning (reporting effects). In a survey of the literature, Kopczuk (2013) concludes that the effects of inheritance taxation on taxable inheritance appear to be relatively small. Goupille-Lebret and Infante (2017) examine changes in the French inheritance taxation and effects on private savings in life insurance funds. Using discontinuities in the tax code with respect to time and age, the authors make an attempt to disentangle real accumulation effects from avoidance responses and find modest effects on real capital accumulation. Kopczuk (2007) made an influential study of estate tax planning in the US, exploiting the receipt of news about terminal illness. The results show that the estates of individuals who received the news substantially decreased in value, primarily due to tax planning.

Another strand of the literature has studied if there are any more general efficiency considerations of inheriting wealth in the context of entrepreneurial activities and family-firm successions. In an unpublished study of Swedish administrative register data on inheritances and firm performance, Escobar (2017) finds that firms whose owners inherit significant amounts tend to survive longer than other, comparable firms do. However, this survival does not seem to be driven by higher productivity, measured as firm profits or owners’ incomes, but instead by enabling small business owners of lower ability to subsist. A similar conclusion regarding the effects of inheritance on entrepreneurial performance was presented by Bennedsen et al. (2007) based on a large Danish micro dataset of family firms, which contains information on whether or not the main owner has deceased. An instrumental variable approach is used exploiting the sex of the first-born child as instrument for whether the firm is taken over by the first-born son in the family or by an external professional CEO. The main result is that there is a large, negative impact of sons inheriting the firm leadership on the firm’s subsequent performance. These studies suggest that inheritances in some cases can be harmful for economic efficiency.

4.5 Corporate taxation

In this paper, we focus on capital taxation at the personal level, and going into detail into how these taxes affect production and firm activity (e.g., the effects of capital taxation on innovation and entrepreneurship or the implications of reduced domestic saving on domestic investment in a world where firms have access to international capital markets) is out of the scope of the present paper. This said, we will briefly discuss a few important aspects related to the taxation of corporate profits. The corporate

52 Gifts transmitted during a person’s life, inter vivos, represent an important part of total lifetime transfers. For this reason, inheritance taxation must be accompanied by gift taxation. Taxing inheritance and inter vivos gifts is commonplace in the industrialized world. A majority of EU’s member currently tax intergenerational transfers, and such taxes exist in a number of large Asian and North American countries.
income tax is a special tax on the profits accruing to private firms. Since all taxes are born by individuals, this also applies to the corporate income tax which formally is a tax on the shareholders of the company which accrues beyond any taxes paid on dividends and capital gains. A classic question in public finance is whether or not the government should tax corporate capital for redistributive reasons because it is concentrated in the upper end of the income distribution. The answer to this question is not clear because a tax on corporate capital may lead to less investment, a lower stock of capital, a higher return to capital, and lower wages. Harberger (1962) found that in a closed economy, a tax a corporate income tax mainly affects the owners of capital, with small effects on wage earners. However, in an open economy, the free mobility of capital changes this result, making it much more likely that wage earners bear a substantial part of its burden. The major constraint on the corporation tax is the possibility for firms to relocate their activities abroad. For this reason, a small open economy must calibrate their corporation tax in accordance with the levels of other similar countries. To identify the extent to which corporate taxation affects wages is a difficult task, both theoretically and empirically, as it represents an exercise in general equilibrium analysis. A recent empirical paper is Fuest, Peichl and Siegloch (2018) which analyzes corporate taxation in Germany and finds that about 40 percent of the burden of the corporate income tax is borne by wage earners.

Three practical arguments in favor of a corporate income tax are often heard. First, and most importantly, the corporate income tax is a complement to the income tax as it is in practice difficult to tax individuals with low labor income and large fortunes derived from inventions, patents or other intellectual property. Incomes from such activities are taxed only at the moment when they accrue to shareholders in the form of dividends or capital gains. In this way, the corporate income tax becomes a way of taxing profits that otherwise would avoid taxation by being kept inside corporations. In addition, in a dual income tax system, the corporate income tax serves to make it less attractive to shift income from the personal to the corporate income tax base.\footnote{An additional reason to tax corporate capital arises if the corporate tax allows to ‘tag’ individuals with high ability. Scheuer (2014) analyzes a model where individuals differ in their income earnings abilities and their cost of setting up a firm. This produces inequality in investment opportunities and therefore the corporate income tax becomes an indirect way of tagging high skill workers to the extent that high skill workers have lower costs of setting up a firm. The overall desirability of corporate income taxation for this purpose depends on how wages are affected.} Second, the corporate income tax is a way to tax foreign investors that do not pay capital income taxes in the host country. The third argument is that the corporation tax can be viewed as a payment for infrastructure that the government provides, such as roads, airports, bankruptcy management, or the value of a stable and secure democracy. The relevance of this argument can however be questioned, as the marginal cost of providing these services are often close to zero and firms contribute to tax revenue through other tax bases by hiring workers.
5. Concluding discussion

The academic economic literature on capital taxation has for decades relied on a set of canonical optimal tax models that forcefully argued that capital taxation has a minor role to play in an optimal tax system. In recent years, however, scholars have increasingly realized that these models do a rather poor job in explaining actual inequality in wealth and capital income, and new theoretical perspectives and empirical observations are challenging the established conventional wisdoms.

In this study, we have asked the following questions: Should capital be taxed? If so, how should it be taxed, considering the many different specific capital tax bases and possible tax schedules? In addressing these questions, we have reviewed both the early and recent theoretical literature on optimal taxation as well as empirical studies on the distribution and aggregate importance of capital and the distortionary costs of capital taxation.

We conclude that capital taxation can be motivated on both equity and efficiency grounds. In particular, the current practice in all developed economies to tax capital income and sometimes wealth appears to be congruent with the recent optimal tax literature. We believe that the most compelling reason to tax capital is that individuals with higher earnings capacities often are those who have high capital income, or have inherited wealth, which makes capital income taxation an efficient complement to progressive labor income taxation.

While progressive capital taxation could be desirable on equity grounds, there are administrative problems with such taxation relating to tax avoidance and tax arbitrage. The Nordic ‘dual’ income tax, which taxes labor income according to a nonlinear progressive tax schedule and capital income according to a proportional rate, could therefore be a constructive way to strike a balance between an optimal and administratively feasible tax system. Even though the dual income tax employs a proportional tax on capital income, and therefore does not reap the potential equity gains of adopting a progressive capital income tax, it is possible to increase the overall progressivity of the tax system by complementing capital income taxes with taxes on property, wealth and inheritance.

Empirical evidence suggests that capital income is more responsive to taxation than labor income, which suggests that the capital income tax rate should be lower than the tax rate on labor income. However, care must be taken so that the divergence is not too large to prevent cross-base income shifting. Further research is needed to quantify the welfare gains of different ways of taxing capital. A quantitative comparison of the relative gains of adopting fully nonlinear non-separable taxes on capital, comprehensive income tax systems, and dual income tax systems, allowing for realistic cross-base shifting is needed to gain further insights into this important policy issue.
Property should be taxed on both equity and efficiency grounds. In particular, the imputed income from owner-occupied housing should be taxed, both because housing provides consumption benefits and because it is an investment with the same properties as a business asset. Liquidity problems should be mitigated by relating property tax payments to household income and employing dampening rules to smooth the changes in tax burden from year to year, or allow for postponement of (some of) the property tax bill until the time the property is sold, but adding a market interest rate to the postponed tax liability to maintain the present value of the tax payment. Capital gains on property should also be taxed, just like any other investment.

Inheritance should probably be taxed. We have discussed the efficiency effects of inheritance taxes, but also that these are notoriously difficult to quantify empirically. It is therefore hard to draw general conclusions about the social desirability of inheritance taxation. However, as inherited wealth does have substantial effects on the wealth distribution and affects the life chances of individuals, we still think a compelling case for inheritance taxation can be made.

An overall ambition should be to tax different investments as uniformly as possible to avoid distortions in investment choices. Maintaining uniformity requires policymakers to withstand pressures from special interest groups as well as solve practical problems associated with the implementation of certain capital taxes (such as liquidity problems).

Finding the right balance between labor and capital taxation is a difficult but important task. It seems appropriate to be equally careful about the distortionary effects of labor income taxes on human capital investments as one should be about the distortionary effects of taxes on physical capital accumulation. Given the large reliance on progressive labor income taxation in many modern welfare states suggests that a better balance between labor and capital taxation could be achieved by increasing the relative tax burden on capital.

There are certain angles and questions that we have not been able to cover. One of these is that governments can facilitate an egalitarian income and wealth distribution in other ways than through taxation, for example by providing individuals better opportunities to create wealth, either through the subsidization of education, reducing transaction costs in credit markets, or lowering labor income taxes. Another question that we have only touched lightly upon is how automation of the workforce and the digitalization of services will affect the importance of capital and capital taxation. Some recent studies argue that automation has resulted in falling wage shares (see, for example, Autor et al. 2017). However, several Western countries have not experienced such declines, despite rising automation. Furthermore, the theoretical basis for understanding the role of capital taxation in light of these changes is still lacking, and we hope that these and related questions will attract considerable attention in the years to come. Finally, we have not analyzed the political economy of capital taxation. From a theory
perspective, Scheuer and Wolitzky (2016) highlight the importance of political constraints by analyzing optimal capital taxation under the threat of a radical political reform that could entail a significant re-distribution of wealth. From an empirical perspective, these questions have been analyzed by Alesina et al. (2018), Fisman et al. (2017), and Bastani and Waldenström (2018). These are very relevant and interesting topics that we also hope will generate much research in the future.
References


Sørensen, P. B. (1994). "From the global income tax to the dual income tax - Recent tax reforms in the Nordic countries", *International Tax and Public Finance* 1(1).


