Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis*

Fatih Guvenen† Burhanettin Kuruscu‡ Serdar Ozkan§

Abstract

Wage inequality has been significantly higher in the United States than in continental European countries (CEU) since the 1970s. Moreover, this inequality gap has further widened during this period as the US has experienced a large increase in wage inequality, whereas the CEU has seen only modest changes. This paper studies the role of labor income tax policies for understanding these facts. We begin by documenting two new empirical facts that link these inequality differences to tax policies. First, we show that countries with more progressive labor income tax schedules have significantly lower *before-tax* wage inequality at different points in time. Second, progressivity is also negatively correlated with the *rise* in wage inequality during this period. We then construct a life cycle model in which individuals decide each period whether to go to school, work, or be unemployed. Individuals can accumulate skills either in school or while working. Wage inequality arises from differences across individuals in their ability to learn new skills as well as from idiosyncratic shocks. Progressive taxation compresses the (after-tax) wage structure, thereby distorting the incentives to accumulate human capital, in turn reducing the cross-sectional dispersion of (before-tax) wages. We find that these policies can account for half of the difference between the US and the CEU in overall wage inequality and 76% of the difference in inequality at the upper end (log 90-50 differential). When this economy experiences skill-biased technological change, progressivity also dampens the rise in wage dispersion over time. The model explains 41% of the difference in the total rise in inequality and 58% of the difference at the upper end.

Keywords: Wage Inequality, Human Capital, Skill-Biased Technical Change, Tax Policies.

*Date: November 8, 2009. For helpful early conversations on this project, we thank Daron Acemoglu, Richard Rogerson, and Tom Sargent. For helpful comments on this draft, we thank Gueorgui Kambourov (our discussant), Tom Krebs, Dirk Krueger, Lars Ljungqvist, Victor Rios-Rull, Gustavo Ventura, Gianluca Violante, and seminar participants at Columbia, Duke, Iowa, NYU, Penn State, Texas A&M, Toronto, Upenn, Universidad Carlos III, Universite de Paris 1-Sorbonne, University of Oslo, Koc, Sabanci; conference participants at the Cowles Foundation Summer Conference at Yale University, Heterogeneous-Agent Models in Macroeconomics Conference at the University of Mannheim, Conference on Consumption and Income Risk at the University of Zurich, the SED conference, and the Midwest Macro Conference. Marina Mendes Tavares provided excellent research assistance. Guvenen acknowledges financial support from the University of Minnesota Grant-in-Aid fellowship. All remaining errors are our own.

†University of Minnesota and NBER; guvenen@umn.edu; http://www.econ.umn.edu/~guvenen

‡UT-Austin; kuruscu@eco.utexas.edu; http://www.eco.utexas.edu/~kuruscu

§University of Pennsylvania; ozkan@econ.upenn.edu; http://www.econ.upenn.edu/~ozkan

1 Introduction

Why is wage inequality significantly higher in the United States (and the United Kingdom) than in continental European countries (CEU)? And why has this inequality gap between the US and the CEU widened substantially since the 1970s (see Table 1)? More broadly, what are the determinants of wage dispersion in modern economies? How do these determinants interact with technological progress and government policies? The goal of this paper is to shed light on these questions by studying the impact of labor market (tax) policies on the determination of wage inequality, using cross-country data.

We begin by documenting two new empirical relationships between wage inequality and tax policy. First, we show that countries with more progressive labor income tax schedules have significantly lower wage inequality at different points in time. The measure of wages we use is "gross before-tax wages"[1] and can therefore be thought of as a proxy for the marginal product of workers. From this perspective, progressivity is associated with a more compressed productivity distribution across workers. Second, we show that countries with more progressive income taxes have also experienced a smaller *rise* in wage inequality over time, and this relationship is especially strong above the median of the wage distribution. This latter finding is intriguing because the substantial part of the rise in wage inequality since 1980 has taken place precisely here—above the median of the distribution (see Table 2). Overall, these findings reveal a close relationship between progressivity and wage inequality, which motivates the focus of this paper. However, on their own, these correlations fall short of providing a quantitative assessment of the importance of the tax structure—e.g., what fraction of cross-country differences in wage inequality can be attributed to tax policies? For this purpose, we build a model.

Specifically, we construct a life cycle model that features some key determinants of wages—most notably, human capital accumulation and idiosyncratic shocks. Here is an overview of the framework. Individuals enter the economy with an initial stock of human capital and are able to accumulate more human capital over the life cycle using a Ben-Porath (1967) style technology (which essentially combines learning ability, time, and existing human capital for production). Individuals can choose to either invest in human capital on the job up to a certain fraction of their time or enroll in school where they can invest full

[1] More precisely, wages are measured before taxes and the employee's social security contributions and also include bonuses and overtime pay when applicable. Therefore, they represent a fairly good measure of the total monetary compensation of a worker.

Table 1: Log Wage Differential Between the 90th and 10th Percentiles

	1978-1982 average	2001-2005 average	Change
Denmark	0.76	0.97	0.20
Finland	0.91	0.89	-0.01
France	1.18	1.08	-0.10
Germany	1.06	1.15	0.09
Netherlands	0.94	1.06	0.12
Sweden	0.71	0.83	0.12
CEU	**0.93**	**1.00**	**0.07**
UK	1.09	1.27	0.18
US	**1.34**	**1.57**	**0.23**

time. We assume that skills are general and labor markets are competitive. As a result, the cost of on-the-job investment will be borne by the workers, and firms will adjust the wage rate downward by the fraction of time invested on the job. Therefore, the cost of human capital investment is the forgone earnings while individuals are learning new skills.

We introduce two main features into this framework. First, we assume that individuals differ in their learning ability. As a result, individuals differ systematically in the amount of investment they undertake and, consequently, in the growth rate of their wages over the life cycle. Thus, a key source of wage inequality in this model is the systematic fanning out of the wage profiles.[2] Second, we allow for endogenous labor supply choice, which amplifies the effect of progressivity, a point that we return to shortly. Finally, for a comprehensive quantitative assessment, we also allow idiosyncratic shocks to workers' labor efficiency, and also model differences in the unemployment insurance and pension systems, which vary greatly across these countries.

The model described here provides a central role for policies that compress the wage structure—such as progressive income taxes—because such policies hamper the incentives for human capital investment. This is because a progressive system reduces *after-tax* wages at the higher end of the wage distribution compared to the lower end. As a result, it reduces the marginal benefit of investment (the higher wages in the future) relative to the marginal cost (the current forgone earnings), thereby depressing investment. A key observation is that this distortion varies *systematically* with the ability level—and, specifically, it worsens

[2]Recent evidence from panel data on individual wages provides support for individual-specific growth rates in wage earnings (cf. Baker (1997), Guvenen (2007, 2009), Huggett, Ventura, and Yaron (2007)).

Table 2: Decomposing the Change in Log 90-10 Wage Differential

	Total Change in Log 90-10	Percentage due to Log 90-50	Log 50-10
CEU	0.07	91%	9%
US	0.23	70%	30%

with higher ability—which then compresses the *before-tax* wage distribution. These effects of progressivity are compounded by endogenous labor supply and differences in average income tax rates: the higher taxes in the CEU reduce labor supply—and, consequently, the benefit of human capital investment—further compressing the wage distribution.

The main quantitative exercise we conduct is the following. We consider the eight countries listed in Table 1, for which we have complete data for all variables of interest. We assume that all countries have the same innate ability distribution but allow each country to differ in the observable dimensions of their labor market structure, such as in labor income (and consumption) tax schedules, and in unemployment insurance and retirement benefits systems. We then calibrate the model-specific parameters to the US data and keep these parameters fixed across countries. The policy differences we consider explain about half of the observed gap in the log 90-10 wage differential between the US and the CEU in the 2000s, and 76% of the wage inequality above the median (log 90-50 differential). When the Frisch elasticity of labor supply is increased to 0.5 from its baseline value of 0.3, the model is able to explain 60% of the log 90-10 differential and virtually all (97% to be exact) of the log 90-50 differential observed in the data. The model explains only about 30% of the difference in the lower tail inequality between the US and the CEU, which is not very surprising, since the human capital mechanism is likely to be more important for higher ability individuals and, therefore, above the median of the distribution. In contrast to the CEU, however, the United Kingdom turns out to be an outlier in the sense that the model is least successful in explaining the features of its wage distribution.

We also provide a decomposition that isolates the roles of (i) the progressivity of income taxes, (ii) average income tax rates, (iii) consumption taxes, and (iv) pension and unemployment insurance systems. We find that progressivity is by far the most important component, accounting for about 68% of the model's explanatory power. As for the remaining three components, each has a similar contribution to the differences in the log 90-10 differential ($\sim$10% each), but consumption taxes are most important for the upper end

(20%) and benefits institutions are most important for the lower end (18%) wage inequality.

A contribution of the present paper that could be of independent interest is the derivation of country-specific effective labor income tax *schedules*, which is, to our knowledge, new to this paper. These schedules are obtained by putting together tax data from different OECD sources and using a flexible functional form that provides a good fit for this relatively diverse set of countries. These tax schedules allow us to measure the progressivity of the (effective) tax structure at different points in the income distribution. This is an essential ingredient in our analysis and could also be useful for studying other questions in the future.

The second question we ask is whether the widening of the inequality gap between the US and the CEU since the late 1970s could also be explained by the same human capital channels discussed earlier. One challenge we face in trying to answer this question is that the tax schedules just described are only available for the years after 2001 (because the detailed information from OECD sources for taxes is only available after that date), whereas the tax structure has changed over time for several of the countries in our sample. Despite this caveat, we cautiously explore how much of the change in the US-CEU inequality gap can be explained with fixed tax schedules.

As shown in Guvenen and Kuruscu (2009), the model described above with the Ben-Porath technology does not have a well-defined notion of returns to skill, which essentially means that changes in the price of human capital (e.g., resulting from skill-biased technical change, SBTC) have no effect on investment behavior. To circumvent this problem, in Section 6, we extend the human capital production technology to a two-factor structure along the lines proposed in that paper.[3] Assuming that all countries have experienced the same degree of SBTC from 1980 to 2003 and using fixed tax schedules over time, the model explains about 41% of the observed gap in the rise in total wage inequality (log 90-10) between the US and the CEU, and about 58% of the difference in the log 90-50 differentials, during this time period. Finally, for two countries in our sample—the US and Germany—we are also able to derive tax schedules for 1983, which reveal significantly more flattening of tax schedules in the US compared to Germany from 1983 to 2003. When these changes in progressivity and SBTC are *jointly* taken into account, the (recalibrated) model generates

[3]Guvenen and Kuruscu (2009) have quantitatively studied a simplified version of this model—one that abstracts from idiosyncratic shocks and endogenous labor supply as well as from all the institutional details studied here—and applied it to the U.S. data. They concluded that even that stark version provides a fairly successful account of several trends observed in the U.S. data since the 1970s. Here, we build on this research by explicitly modeling labor market institutions and allowing for idiosyncratic shocks and endogenous labor supply to understand the role of tax policy for wage inequality.

a much larger rise in inequality in the US than in Germany and, in fact, overestimates the actual widening of the inequality gap between these countries by 16%.

Overall, these results illustrate how government policies can strongly influence the response of an economy to technological change by distorting individuals' incentives to undertake human capital investment, which keeps inequality low but at the cost of lower aggregate output. To highlight this point, we briefly discuss the implications of the model for some macroeconomic variables, such as GDP, hours, unemployment rates, and so on.

Related Literature. Some previous papers have also examined the US-CEU differences in wage inequality, although using techniques that are quite different from the present paper; see, e.g., Blau and Kahn (1996), Kahn (2000), and Gottschalk and Joyce (1998). These papers mainly use regression analyses and conclude that unionization, centralized bargaining, and minimum wage laws are important for understanding European wage inequality data. An important point to note, however, is that these studies do not consider the role of progressive taxation in their regression analyses. Because countries with more rigid institutions also have more progressive tax systems (see Appendix A), this omission could attribute the effect of progressivity to these other institutions. A notable exception in this literature is Acemoglu (2003), who constructs a fully specified model in which wage compressing institutions in the CEU affect the incentives of firms such that they adopt technologies that are less skill biased than in the US. Thus, in his model inequality rises less in Europe because the rise in skill demand is slower in that region. His paper highlights a novel channel, which can be complementary to the mechanism studied in this paper.

In terms of methodology, this paper is also related to the recent macroeconomics literature that has written fully specified models to address US-CEU differences in labor market outcomes. Prominent examples include Ljungqvist and Sargent (1998), Ljungqvist and Sargent (2008), and Hornstein, Krusell, and Violante (2007), who focus on unemployment rates, and Prescott (2004), Ohanian, Raffo, and Rogerson (2006), and Rogerson (2008), who study labor hours differences. Several of these papers rely on representative agent models and are, therefore, silent on wage inequality; and those that do allow for individual-level heterogeneity do not address differences in wage inequality.[4] In terms of modeling choices,

[4] A notable exception is Hornstein, Krusell, and Violante (2007), who do study the implications of their framework for wage inequality but conclude that it does not generate much wage dispersion or differences in wage inequality across countries, despite having successful implications for unemployment rates and the labor share.

the closest framework to ours is Kitao, Ljungqvist, and Sargent (2008), who study a rich life cycle framework with human capital accumulation and job search, and model the benefits system. Their goal is to explain the different unemployment patterns over the *life cycle* in the US and Europe.

Finally, a number of recent papers share some common modeling elements with ours but address different questions. Important examples include Krebs (2003), Caucutt, Imrohoroglu, and Kumar (2006), Huggett, Ventura, and Yaron (2007), and Erosa and Koreshkova (2007). Krebs (2003) studies the impact of idiosyncratic shocks on human capital investment and shows that reducing income risk can increase growth, in contrast to the standard incomplete markets literature, which typically reaches the opposite conclusion. Caucutt, Imrohoroglu, and Kumar (2006) develop an endogenous growth model with heterogeneity in income. They show that a reduction in the progressivity of tax rates can have positive growth effects even in situations where changes in flat rate taxes have no effect. Another important contribution is Huggett, Ventura, and Yaron (2007), who study the distributional implications of the Ben-Porath model and estimate the sources of lifetime inequality using US earnings data. Finally, the interaction of human capital investment and progressive taxes is also present in Erosa and Koreshkova (2007), who investigate the effects of replacing the current U.S. progressive income tax system with a proportional one in a dynastic model. They find a large positive effect on steady state output, which comes at the expense of higher inequality. Although our paper has many useful points of contact with this body of work, to our knowledge, the combination of human capital accumulation, ability heterogeneity, progressive taxation, and endogenous labor supply is new to this paper, as is the attempt to explain cross-country inequality facts in such a framework.

The next section starts with a stylized model to explain the various channels through which tax policy affects wage inequality. It then explains how the country-specific tax schedules are estimated and uses the estimates to document some empirical links between taxes and inequality. Sections 3 and 4 describe the main model and the parametrization. Section 5 presents the cross-sectional quantitative results and sensitivity analyses. Section 6 extends this model to examine the evolution of inequality over time. Section 7 concludes.

2 US versus CEU: Differences in Empirical Trends

In this section, we document two new empirical relations between wage inequality and the progressivity of the tax policy. To this end, we begin with a stylized version of the more general model studied in Section 3 that illustrates the key mechanisms at work and will allow us to define different measures of progressivity subsequently used in documenting the empirical facts. We then discuss how the tax schedules are derived for each country and present the empirical findings in Section 2.3.

2.1 Model 0: Intuition in a Stylized Framework

Consider an individual who derives utility from consumption and leisure and has access to borrowing and saving at a constant interest rate, r. Let β be the subjective time discount factor and assume $\beta(1+r)=1$. Each period individuals have one unit of time endowment that they allocate between leisure and work ($n \in [0,1]$). While working, individuals can accumulate new human capital, Q, according to a Ben-Porath style technology. Specifically, $Q = A^j\,(hin)^\alpha$, where h denotes the individual's current human capital stock, i denotes the fraction of working time (n) spent learning new skills, and A^j is the learning ability of individual type j. We assume that skills are general and labor markets are competitive. As a result, the cost of human capital investment is completely borne by workers, and firms adjust the hourly wage rate downward by the fraction of time invested on the job (equation (2)). Finally, labor earnings are taxed at a rate given by the average tax function $\bar{\tau}_n(y)$, and the marginal tax rate function is denoted by $\tau(y)$. The problem of a type j individual can be written as

$$\max_{c_s,a_{s+1},i_s} \sum_{s=1}^{S} \beta^{s-1} u(c_s, 1-n_s)$$

$$\text{s.t.} \qquad c_s + a_{s+1} = (1-\bar{\tau}_n(y_s))y_s + (1+r)a_s$$

$$h_{s+1} = h_s + A^j\,(h_s i_s n_s)^\alpha \tag{1}$$

$$y_s = P_H h_s (1-i_s) n_s \tag{2}$$

Using the fact that $Q_s^j = A^j\,(h_s i_s n_s)^\alpha$, the opportunity "cost of investment" (i.e., $h_s i_s n_s$) can be written as $C_j(Q_s^j) = (Q_s^j/A^j)^{(1/\alpha)}$, which will play a key role in the optimality conditions that follow. Now, it is useful to distinguish between two cases.

Inelastic Labor Supply. First, suppose that labor supply is inelastic. The optimality condition for human capital investment is (assuming an interior solution)

$$(1-\tau(y_s))\,C_j'(Q_s^j) = \{\beta(1-\tau(y_{s+1})) + \beta^2\,(1-\tau(y_{s+2})) + ... + \beta^{S-s}\,(1-\tau(y_S))\}. \tag{3}$$

The left-hand side is the marginal cost of investment and the right-hand side is the marginal benefit, which is given by the present discounted value of net wages in all future dates earned by the extra unit of human capital. Notice that both the marginal cost and benefit of investment take into account the marginal tax rate faced by the individual. To understand the effect of taxes, first consider the case where taxes are flat rate, i.e., $\tau'(y) \equiv 0$. In this case, all terms involving taxes cancel out and the first order condition reduces to

$$C_j'(Q_s^j) = \{\beta + \beta^2 + ... + \beta^{S-s}\}.$$

Thus, flat-rate taxes have *no effect* on human capital investment. This is a well-understood insight that goes back to at least Heckman (1976) and Boskin (1977).[5]

Now consider progressive taxes, i.e., $\tau'(y) > 0$. We rearrange equation (3) to get:

$$C_j'(Q_s^j) = \{\beta \underbrace{\frac{1-\tau(y_{s+1})}{1-\tau(y_s)}}_{\text{Progressivity Discount}} + \beta^2 \frac{1-\tau(y_{s+2})}{1-\tau(y_s)} + ... + \beta^{S-s} \underbrace{\frac{1-\tau(y_S)}{1-\tau(y_s)}}_{\text{Progressivity Discount}}\}. \tag{4}$$

As long as the individual's earnings grow over the life cycle and the tax structure is progressive, all tax ratios on the right-hand side will be smaller than one, which will depress the marginal benefit of investment and in turn dampen human capital accumulation. Thus, each of these tax ratios captures the *progressivity discount* that effectively reduces the value of higher wage earnings in the future when compared to the lower forgone wage earnings today. To draw an analogy to the taxation literature (c.f. Prescott (2004), Ohanian, Raffo, and Rogerson (2006), etc.), it is useful to focus on a closely related measure—what we refer to as the *progressivity wedge*—which is essentially one minus the progressivity discount:

$$PW(y_s, y_{s+k}) \equiv 1 - \frac{1-\tau(y_{s+k})}{1-\tau(y_s)} = \frac{\tau(y_{s+k})-\tau(y_s)}{1-\tau(y_s)}. \tag{5}$$

[5] With pecuniary costs of investment, flat taxes can affect human capital investment, as shown by King and Rebelo (1990) and Rebelo (1991). Similarly, Lucas (1990) shows that flat taxes can have a negative impact on human capital investment when labor supply is elastic.

These wedges provide a key measure of the distortion created by progressive taxes. A progressivity wedge of zero corresponds to flat taxes, and the distortion grows with the wedge. To understand the effect of progressive taxes on wage inequality, first note that the distortion created by progressivity differs systematically across ability levels. At the low end, individuals with very low ability whose optimal plan involves no human capital investment in the absence of taxes would experience no wage growth over the life cycle and, therefore, no distortion from progressive taxation. At the top end, individuals with high ability whose optimal plan implies low wage earnings early in life and very high earnings later face very large wedges, which depress their investment. Thus, progressivity reduces the cross-sectional dispersion of human capital and, consequently, the wage inequality in an economy, even with inelastic labor supply.[6]

Endogenous Labor Supply. Second, consider now the the case with elastic labor supply. The first order condition can be shown to be as follows:

$$C_j'(Q_s^j) = \{\beta \frac{1-\tau(y_{s+1})}{1-\tau(y_s)} n_{s+1} + \beta^2 \frac{1-\tau(y_{s+2})}{1-\tau(y_s)} n_{s+2} + ... + \beta^{S-s} \frac{1-\tau(y_S)}{1-\tau(y_s)} n_S\}, \tag{6}$$

where now the marginal benefit accounts for the utilization rate of human capital, which depends on the labor supply choice (for derivation, see Appendix B.1). Now, once again, consider the effect of flat rate taxes. The intratemporal optimality condition implies that labor supply depends negatively on the tax rate and positively on the level of human capital. A higher tax rate depresses labor supply choice (as long as the income effect is not too large), which then reduces the marginal benefit of human capital investment, which reduces the optimal level of human capital. But labor supply in turn depends on the level of human capital, which further depresses labor supply, the level of human capital and so on. Therefore, with endogenous labor supply, even a flat-rate tax has an effect on human capital investment, which can also be large because of the amplification described here.[7]

Because average labor hours differ significantly across countries and over time (c.f.,

[6] It is easy to see that in a model with retirement (as in the next section), a redistributive pension system will have an effect that would work very similarly to progressive income taxation. The same is true for the unemployment insurance system, which dampens the incentives to invest, although this is likely to be more important at the lower end of the income distribution. We incorporate both into the full model later.

[7] Similarly, policies that restrict labor supply (such as the 35-hour workweek law implemented in France during much of the 2000s) will also depress human capital accumulation and compress the wage distribution. This illustrates a situation where unions (who lobbied for the restrictions imposed in France) can affect even inequality at the upper end.

Prescott (2004), Ohanian, Raffo, and Rogerson (2006)), it is also useful to consider a second measure of wedge that takes into account each country's utilization rate of its human capital (relative to the average country in the sample) in addition to its tax structure. Formally, for country i, what we now call the *progressivity wedge**, is defined as

$$PW_i^*(y_s, y_{s+k}) = 1 - \frac{1-\tau(y_{s+k})}{1-\tau(y_s)}\left(\frac{n_i}{n_{ALL}}\right), \tag{7}$$

where n_i is the hours per person in country i and, similarly, n_{ALL} is the average of hours across all countries in the sample.[8]

In summary, the stylized model studied here implies that countries with more progressive tax systems will have a lower wage inequality. As will become clear later, these countries will also experience smaller rises in wage inequality in response to SBTC.

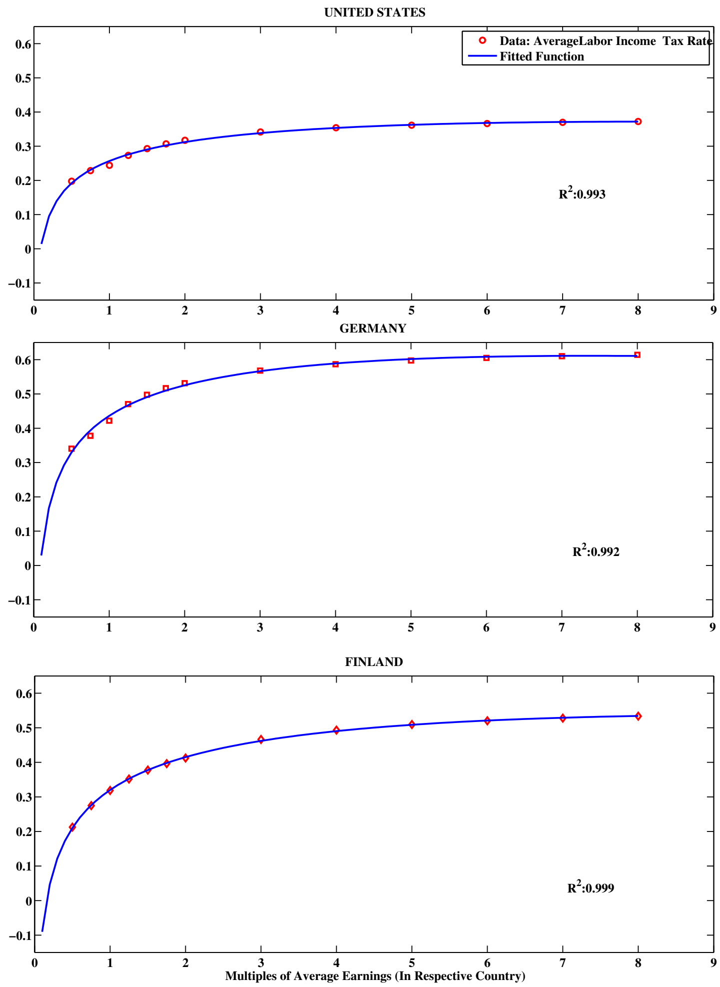
2.2 Deriving Country-Specific Tax Schedules

For each country, we follow the same procedure described here. First, the OECD tax database provides a calculator that estimates the total labor income tax for all income levels between half of average wage earnings (hereafter, AW) to two times AW. The calculation takes into account several types of taxes (central government, local and state, social security contributions made by the employee, and so on), as well as many types of deductions and cash benefits (dependent exemptions, deductions for taxes paid, social assistance, housing assistance, in-work benefits, etc.).[9] Using this tool, we calculate the average labor income tax rate, $\bar{\tau}(y)$, for 50%, 75%, 100%, 125%, 150%, 175%, and 200% of AW. However, tax rates beyond 200% of AW are also relevant when individuals solve their dynamic program. Fortunately, another piece of information is available from the OECD: specifically, we also have the *top* marginal tax rate and the *top* bracket corresponding to it for each country. As described in more detail in Appendix C.1, we use this information to generate average tax rates at income levels beyond two times AW. Then, we fit the following smooth function to

[8]Notice that because of the rescaling by n_{ALL}, if a country has sufficiently high labor hours and low progressivity, this wedge measure can become negative (e.g., the US). Therefore, this new measure is defined relative to a given sample of countries, but is still informative about the relative return to human capital within a group of countries, which is the focus of this paper.

[9]Non-wage income taxes (e.g., dividend income, property income, capital gains, interest earnings) and non-cash benefits (free school meals or free health care) are not included in this calculation.

Figure 1: Estimated Average Tax Rate Functions, Selected OECD Countries, 2003



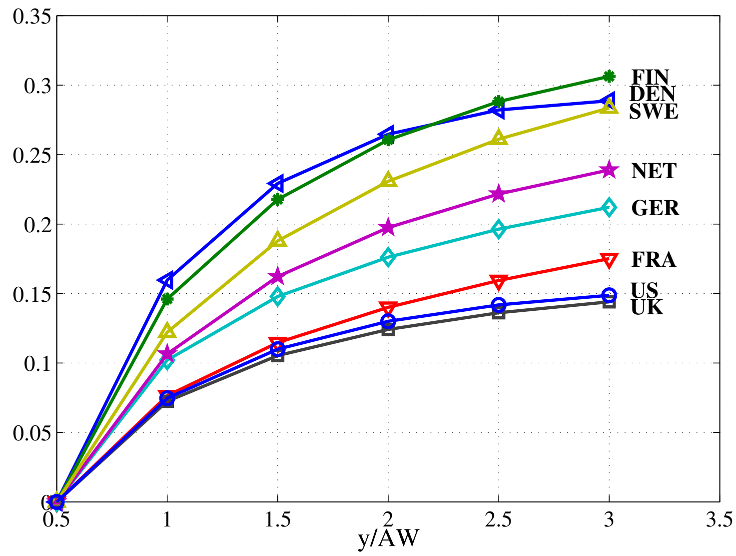
the available data points:[10]

$$\bar{\tau}(y/AW) = a_0 + a_1(y/AW) + a_2(y/AW)^{\phi}. \tag{8}$$

The parameters of the estimated average tax functions for all countries are reported in Appendix A (Table A.2), along with the R^2 values. Although the assumed functional form allows for various possibilities, all fitted tax schedules turn out to be increasing and concave. The lowest R^2 is 0.984 and the mean is 0.991, indicating a fairly good fit. In Figure 1 we plot

[10]We have also experimented with several other functional forms, including a popular specification proposed by Guoveia and Strauss (1994), commonly used in the quantitative public finance literature (cf. Castañeda, Díaz-Giménez, and Ríos-Rull (2003), Conesa and Krueger (2006), and the references therein). However, we found that the functional form used here to provides the best fit across the board for these relatively diverse set of countries, as seen from the high R^2 values in Table A.2.

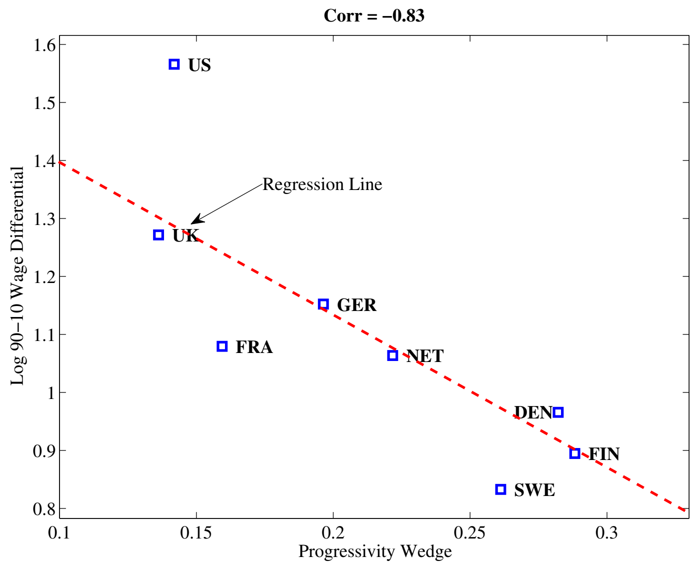
Figure 2: Progressivity Wedges At Different Income Levels: $1 - \frac{1-\tau(k \times 0.5)}{1-\tau(0.5)}$ for $k = 2, 3, .., 6$.



the estimated functions for three countries: one of the two least progressive (United States), the most progressive (Finland), and one with intermediate progressivity (Germany).

Figure 2 plots the progressivity wedges for the eight countries in our sample. Specifically, each line plots $PW(0.5, 0.5k)$ for $k = 1, 2, ..., 6$, which are essentially the wedges faced by an individual who starts life at half the average earnings in that country and looks toward an eventual wage level that is up to six times his initial wage. As seen in the figure, countries are ranked in terms of their progressivity, consistent with one could conjecture: the US and the UK have the least progressive tax system, whereas Scandinavian countries have the most progressive one, with larger continental European countries scattered between these two extremes. The differences also appear quantitatively large (although a more precise evaluation needs to await the full-blown model in Section 3): for example, the marginal benefit of investment for a young worker who invests today when her wage is $0.5 \times AW$ and aims to earn $2 \times AW$ in the future is 13% lower than a flat-tax system in the US and the UK, compared to 27% lower in Denmark and Finland. These differences grow with the ambition level of the individual, dampening human capital investment, especially at the top of the distribution.

Figure 3: **Progressivity Wedge and the Log 90-10 Wage Dispersion in 2003**. A wedge of zero corresponds to flat taxation (no distortion), and progressivity increases along the horizontal axis. The wedge measure used corresponds to *PW(0.5,2.5)* as defined in the text.



2.3 Taxes and Inequality: Cross-Country Empirical Facts

As explained earlier, the average labor income tax schedule in 2003 has been estimated for each of the eight countries listed in Table 1. Using these schedules, we normalize AW in each country to 1 and focus on the progressivity wedge between half the average earnings and 2.5 times the average earnings: $PW(0.5, 2.5)$. Similarly, when we calculate PW^* for a given country, we use the average hours per person in that country between 2001 and 2005 for n_i in equation (7), and the average of the same variable across all countries for n_{ALL}.

The wage inequality data come from the OECD's Labour Force Survey database and are derived from the gross (i.e., before-tax) wages of full-time, full-year (or equivalent) workers.[11] This is the appropriate measure for the purposes of this paper, as it more closely corresponds to the marginal product of each worker (and, hence, her wage) in the model.

[11] The definition of gross wages is given in footnote 1. An exception to this definition is France, for which wage earnings are *net* of employee social security contributions. Also, in contrast to the other countries in the sample, France excludes "agricultural and general government workers and household service workers" from its samples when reporting wage data. Despite these caveats, we are including France in our sample because it is not clear how much these differences affect the final wage inequality numbers. To get an idea, we have compared the wage inequality figures from our main data to another source for France, also provided in the OECD Labour Force Survey (reported as the *GAE0* variable), which includes all workers and reports gross wages but is only available from 2002 to 2005. At least during this period, the two data sources agree extremely well, which is reassuring.

Table 3: Cross-Correlation of $PW(k,m)$ and Log 90-10 Wage Differential

$k \downarrow$	$m \rightarrow$	1.0	1.5	2.0	2.5	3.0
			1980			
0.5		0.82	0.85	0.86	0.87	0.88
1.0			0.87	0.87	0.87	0.86
1.5				0.86	0.84	0.79
2.0					0.76	0.65
			2003			
0.5		0.73	0.77	0.80	0.83	0.85
1.0			0.81	0.84	0.87	0.89
1.5				0.87	0.88	0.88
2.0					0.87	0.82

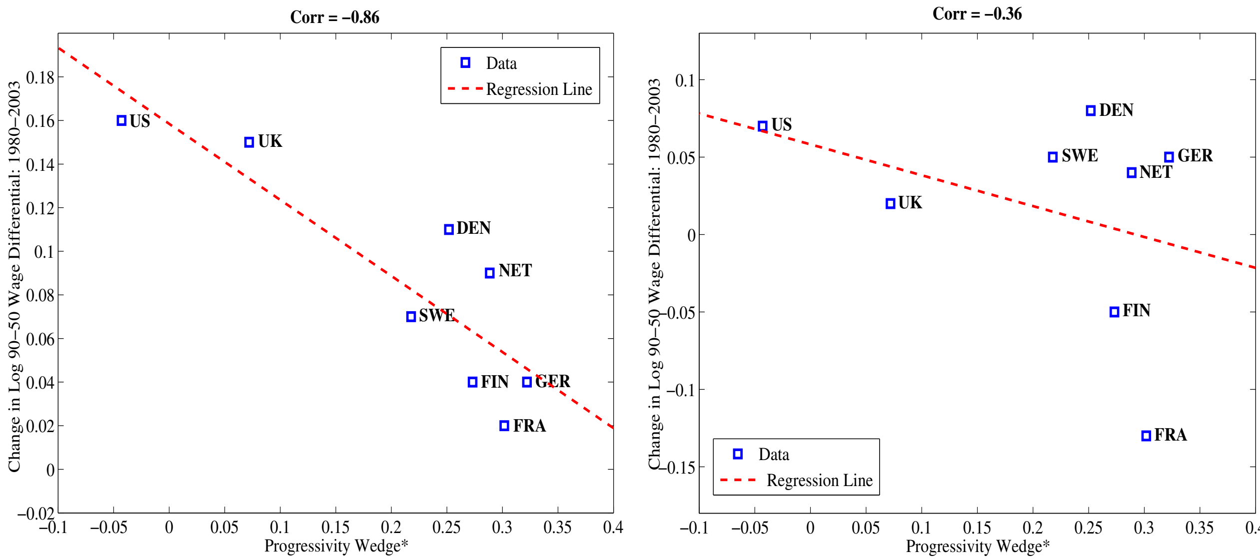
The fact that the inequality data pertain to before-tax wages is important to keep in mind; if the data were after-tax wages, the correlation between the progressivity of taxes and inequality would be mechanical and, thus, not surprising at all.

Figure 3 plots the relationship between the log 90-10 wage differential and progressivity wedge in the 2000s. Countries with a smaller wedge—meaning a less progressive tax system and, therefore, smaller distortion in human capital investment—have a higher wage inequality. The relationship is also quite strong with a correlation of 0.83. Repeating the same calculation using the progressivity wedge* yields a correlation of 0.75. Both relationships are consistent with the human capital model with progressive taxes presented earlier. Moreover, this strong relationship is robust to using wedges calculated from different parts of the wage distribution. This is seen in Table 3, which reports the correlation between the log 90-10 wage differential and $PW(k,m)$ as k and m are varied over a wide range.[12]

We next turn to the change in inequality over time. Figure 4 plots the progressivity wedge* versus the change in the log 90-50 (left panel) and the log 50-10 (right panel) wage differentials. Countries with a more progressive tax system in the 2000s have experienced a smaller rise in wage inequality since the 1980s. The relationship is especially strong at the top of the wage distribution and weaker at the bottom: the correlation between progressivity and the change in the 90-50 differential is very strong (-0.86), whereas the correlation with

[12]The same table also reports the correlation for each country in 1980, even though the wedges are still the ones obtained using the 2003 tax schedules. Surprisingly, even in this case, the correlation is as strong as before. One possible explanation is that the relative ranking of inequality across these countries might not have changed much since 1980. Indeed, the correlation between the log 90-10 wage differential in 2003 and 1980 is 0.87.

Figure 4: **Progressivity Wedge* and Change in Log 90-50 (Left) and 50-10 (Right) Differentials: 1980 to 2003.** The wedge measure for country i is $PW_i^*(0.5, 2.5)$ as defined in equation (7). Progressivity increases along the horizontal axis.



the 50-10 differential is much weaker (only −0.36); see Figure 4). This result is consistent with the idea that the distortion created by progressivity is likely to be felt especially strongly at the upper end where human capital accumulation is an important source of wage inequality, but less so at the lower end, where other factors, such as unionization, minimum wage laws, and so on, could be more important.

Finally, Table 4 gives a more complete picture of the differences between the two definitions of wedges. The top panel reports the correlation of each wedge measure with log wage differentials, which reveals that the adjustment for utilization rates through labor hours makes little difference in the correlations in 2003 but has a somewhat larger effect (reduction) in the correlations in 1980. However, with either measure, progressivity is negatively correlated with inequality even when one focuses on different parts of the distribution. This picture changes when we turn to the change in inequality over time (bottom panel). Now the simple wedge measure has a rather low correlation with log wage differentials (the strongest is with log 90-50 and that is −0.39). However, adjusting for hours per person increases these correlations significantly to −0.63 for the log 90-10 differential, and to −0.86 for the log 90-50 differential (which is plotted in the left panel of Figure 4). We conclude that the relationship between the wedge measures and cross-sectional inequality is quite robust, whereas the change in inequality over time is more sensitive to the adjustment by hours per

Table 4: Correlation Between Progressivity Measures and Wage Dispersion

	Measure of Wedge:	
	$PW(0.5, 2.5)$	$PW^*(0.5, 2.5)$
Log wage differentials	Year: 2003	
90-10	$-.83$	$-.75$
90-50	$-.84$	$-.73$
50-10	$-.73$	$-.67$
	Year: 1980	
90-10	$-.87$	$-.51$
90-50	$-.74$	$-.38$
50-10	$-.88$	$-.57$
	Change from 1980 to 2003	
90-10	$-.11$	$-.63$
90-50	$-.39$	$-.86$
50-10	$.14$	$-.36$

person. Since our full model includes a labor supply choice, this latter measure will become the more relevant one, as we shall see in the next section.

3 Model 1: For Cross-Sectional Analysis

The model we use for the cross-sectional analysis is a richer version of the basic framework presented in Section 2.1. Each individual has one unit of time in each period, which she can allocate to three different uses: work, leisure, and human capital investment. Preferences over consumption, c, and leisure time, $1-n$, are given by this common separable form:

$$u(c,n) = \log(c) + \psi \frac{(1-n)^{1-\varphi}}{1-\varphi}. \tag{9}$$

If an individual chooses to work, as before, she can allocate a fraction (i) of her working hours (n) to human capital investment. However, more realistically, we now assume that $i \in [0,\chi]$, where $\chi < 1$. An upper bound less than 100% on on-the-job investment can arise, for example, because the firm incurs fixed costs for employing each worker (administrative burden, cost of office space, etc.) or as a result of minimum wage laws. Individuals *can* invest full-time by attending school ($i = 1$) and enjoy leisure for the rest of the time. Thus, the choice set is $i \in [0,\chi] \cup \{1\}$, which is non-convex when $\chi < 1$. Finally, human capital depreciates every period at rate $\delta < 1$. Except for the differences described here, the human capital accumulation process is the same as the stylized model described in Section 2.1.

As before, human capital is produced according to a Ben-Porath technology: $Q = A^j (hin)^\alpha$. A key parameter in this specification is A^j, which determines the productivity of learning. The heterogeneity in A^j implies that individuals will differ systematically in the amount of human capital they accumulate and, consequently, in the growth rate of their wages over the life cycle. This systematic fanning out of wage profiles is the major source of wage inequality in this model. Also, as can be seen from the optimality conditions (for example, (6)), the price of human capital has no effect in this model other being a scaling factor. Thus, for simplicity we set $P_H = 1$ in the rest of the cross-sectional analysis.

An individual may choose to be unemployed at age s, $n_s = 0$, in which case she receives unemployment benefit payments as specified later. Individuals retire at age R and receive constant pension payments every year until they die at age T. The benefits system is described in more detail later on.

Idiosyncratic Shocks and Earnings. Individuals receive idiosyncratic shocks to the efficiency of the labor they supply in the market. Specifically, when an individual devotes $n_s(1 - i_s)$ hours producing for his employer, his effective labor supply becomes $\epsilon n_s(1 - i_s)$, where the ϵ shocks are generated by a stationary Markov transition matrix $\Pi(\epsilon' \mid \epsilon)$ that is identical across agents and over the life cycle. The observed *total wage income* of an individual is $y_s^j \equiv \epsilon h_s^j n_s^j (1 - i_s^j)$, and the *hourly wage* rate is simply $w_s^j = y_s^j / n_s^j$.

3.1 Government: Taxes and Transfers

Unemployment and Pension Benefits. The unemployment benefit system is modeled so as to capture the salient features of each country's actual system in a relatively parsimonious manner. For computational reasons, we make some simplifying assumptions to the actual systems implemented by each country. Specifically, if a worker becomes unemployed at age s, the initial level of the unemployment payment she receives is an increasing function of her years of work before becoming unemployed, denoted by m, and also (typically) decreases with the duration of the unemployment spell. Furthermore, in most countries the replacement *rate* falls with the level of pre-unemployment income, which is also partly captured here. Let $\Phi(y^*, m, s)$ denote the unemployment benefit function of an s-year-old individual with m years of employment before becoming unemployed. Although, in reality, unemployment payments depend on the pre-unemployment earnings, y_{s-1}, making this dependence explicit will add an additional state variable into an already demanding non-convex computational problem. Thus, we simplify the problem by assuming that Φ instead depends on y^*, which is the income the individual would have earned in the current state at age s if he did not have the option of receiving unemployment insurance. For the precise mathematical problem that yields y^*, see Appendix B.2.

After retirement individuals receive constant pension payments every period. Essentially, the pension of a worker with ability level j depends on the *average lifetime earnings* of workers with the same ability level (denoted by $\overline{y}^j$) as well as on the number of years the worker has been employed up to the retirement age (denoted by m^R) subject to a maximum years of contribution, $\overline{m}$. The pension function is denoted as $\Omega(\overline{y}^j, m^R)$.[13]

The Tax System and the Government Budget. The government imposes a flat-rate consumption tax, $\bar{\tau}_c$, as well as a potentially progressive labor income tax, $\bar{\tau}_n(y)$.[14] The collected revenues are used for two main purposes: (i) to finance the benefits system, and (ii) to finance government expenditure, G, that does not yield any direct utility to consumers (because of either to corruption or waste). The residual budget surplus or deficit, Tr, is distributed in a lump-sum fashion to all households regardless of employment status.

3.2 Individuals' Dynamic Program

Individuals are able to trade a full set of one-period Arrow securities. A security that promises to deliver one unit of consumption good in state ϵ' in the next period costs $q(\epsilon'|\epsilon)$ in state ϵ today. Let I_n be an indicator that is equal to 0 if the agent is unemployed and 1 otherwise (if a worker or student). The dynamic program of a typical individual is given by

$$\begin{aligned}
V(h, a, m; \epsilon, s) &= \max_{c,n,i,a'(\epsilon')} \left[u(c, n) + \beta E\left(V(h', a'(\epsilon'), m'; \epsilon', s+1)|\epsilon\right)\right] && (10)\\
\text{s.t.} & \\
(1+\bar{\tau}_c)c + \sum q(\epsilon' \mid \epsilon)a'(\epsilon') &= (1-\bar{\tau}_n(y))y + a + Tr, && (11)\\
y &= I_n \times \epsilon h(1-i)n + (1-I_n) \times \Phi(y^*, m, s), && (12)\\
h' &= (1-\delta)h + A(hni)^\alpha, \qquad l' = (1-\delta)l, && (13)\\
m' &= m + I_n \times 1\{i < 1\}, && (14)\\
i &\in [0, \chi] \cup \{1\},
\end{aligned}$$

[13]In reality, pension payments depend on the workers' own earnings history, but modeling this explicitly also adds an extra state variable, which this simplified structure avoids.

[14]Because capital is mobile internationally, it is harder to justify using country-specific tax rates on capital income, unlike for labor and consumption, which are almost always taxed at destination (or the country of residence of the worker). In particular, the mobility of capital implies the equalization of after-tax rates across countries of comparable assets. For these reasons, we abstract from capital income taxes.

where we suppress ability type for clarity. Notice from equation (13) that individuals cannot accumulate human capital while unemployed ($n = 0$). Of course, an individual may return to school after losing her job, in which case she is considered a student and not unemployed. Finally, equation (14) makes clear that m' increases only when agents work and not when they are enrolled in school (i.e., $i = 1$).

After retirement, individuals receive a pension and there is no human capital investment. Since there is no uncertainty during retirement, a riskless bond is sufficient for smoothing consumption. Therefore, the problem of a retired agent at age $s > R$ can be written as

$$\begin{aligned} W^R(a, \overline{y}^j, m^R; s) &= \max_{c,a'} \left[u(c, \bar{n}) + \beta W^R(a', \overline{y}^j, m^R; s+1) \right] \qquad (15) \\ \text{s.t} \qquad (1 + \bar{\tau}_c)c + qa' &= (1 - \bar{\tau}_n(y_s))y_s + a + Tr \\ y_s &= \Omega(\overline{y}^j, m^R). \end{aligned}$$

Definition 1 *A stationary competitive equilibrium for this economy is a set of equilibrium decision rules, $c(x)$, $n(x)$, $Q(x)$, $i(x)$, and $a'(\epsilon', x)$; value functions, $V(x)$ and $W^R(x)$, for working and retirement periods, respectively, where $x = (h, a, m; \epsilon, s, j)$ (notice the inclusion of j into this vector); a pricing function for Arrow securities, $q(\epsilon'|\epsilon)$, and a measure $\Lambda(x)$ such that*

1. *Given the labor income tax function, $\bar{\tau}(y)$, consumption tax, $\bar{\tau}_c$, transfers, Tr, and government policy functions, Φ and Ω, individuals' decision rules and value functions solve problems in (10) to (14) and in (15).*

2. *Asset markets clear: $\int_{x(:,\epsilon=\tilde{\epsilon})} a'(\epsilon', x) d\Lambda(x) = 0$ for all combinations of $(\tilde{\epsilon},\epsilon')$.*[15]

3. *$\Lambda(x)$ is generated by individuals' optimal choices.*

4. *The government budget balances:*

$$\int_{x(:,s<R)} \bar{\tau}_n(y(x))y(x)d\Lambda(x) + \int_x \bar{\tau}_c c(x)d\Lambda(x) = G + Tr$$
$$+ \int_{x(:,s<R)} \Phi(y^*(x), m, s)I(n(x) = 0)d\Lambda(x) + \sum_{s=R}^{T} \int_{x(:,s=R-1)} \Omega(\overline{y}^j, m^R(x))d\Lambda(x).$$

[15]The notation $x(:, \epsilon = \tilde{\epsilon})$ indicates that the integral is taken over the entire domain of variables in state vector x, except for ϵ, which is set equal to $\tilde{\epsilon}$. Others below are defined analogously.

The first term in the government's budget is the total tax revenue from labor income collected from all agents who are working and younger than retirement age. Similarly, the second term is the total tax revenue from the consumption tax, but it is collected from all agents including the retirees. On the right-hand side, the pension payments only depend on a worker's ability through $\overline{y}^j$ and the number of years she worked until retirement ($m^R(x)$), which in turn depends on the full state vector x at age $R-1$. Therefore, we integrate the pension payments over the full state vector x conditioning on age $R-1$ and then sum the same amount over all ages greater than $R-1$ to find total pension payments.

4 Quantitative Analysis

In this section, we begin by discussing the parameter choices for the model. Our basic calibration strategy is to take the United States as a benchmark and pin down a number of parameter values by matching certain targets in the US data.[16] We then assume that other countries share the same parameter values with the US along unobservable dimensions (such as the distribution of learning ability), but differ in the dimensions of their labor market policies that are feasible to model and calibrate (specifically, consumption and labor income tax schedules, the retirement pension system, and the unemployment insurance system). We then examine the differences in economic outcomes—specifically in wage dispersion, output, and labor supply—that are generated by these policy differences alone.

4.1 Calibration

A model period corresponds to one year of calendar time. Individuals enter the economy at age 20 and retire at 65 ($S = 45$). Retirement lasts for 20 years and everybody dies at age 85. The net interest rate, r, is set equal to 2%, and the subjective time discount rate is set to $\beta = 1/(1+r)$.[17] The curvature of the human capital accumulation function, α, is set equal to 0.80, broadly consistent with the existing empirical evidence, and the maximum investment allowed on the job, χ, is set to 0.50 (see Guvenen and Kuruscu (2009) for further justification of these parameter choices).

[16]Taking the US as the benchmark is motivated by the fact that its economy is subject to much less of the labor market rigidities present in the CEU—such as unionization and other distorting institutions. Because these institutions are not modeled in this paper, the US provides a better laboratory for determining the unobservable parameters than other countries where these distortions could be more important for wage determination.

[17]This interest rate should be thought of as the "after-tax" rate, since we do not model taxes on savings explicitly.

Utility Function. The utility function given in (9) has two parameters to calibrate: the curvature of leisure, φ, and the utility weight attached to leisure, ψ. These parameters are jointly chosen to pin down the average hours worked in the economy, as well as the average Frisch labor supply elasticity. We assume that each individual has 100 hours of discretionary time per week (about 14 hours a day), and taking 40 hours per week as the average labor supply for employed workers in the US implies $\overline{n} = 0.4$. With power utility, the theoretical Frisch elasticity of labor is equal to $\frac{1-n}{n}\frac{1}{\varphi}$. Because of heterogeneity across individuals, labor supply varies in the population, so there is a distribution of Frisch elasticities. We simply target the Frisch elasticity implied by the average labor hours, $\overline{n}$. The empirical target we choose is 0.3, which is consistent with the estimates surveyed by Browning, Hansen, and Heckman (1999), which range from zero to 0.5. Although it is common to use higher elasticity values in representative agent macro studies (e.g., Prescott (2004) among many others), values of 0.5 or lower are more common in quantitative models with heterogeneous agents (cf. Heathcote, Storesletten, and Violante (2008), Erosa, Fuster, and Kambourov (2009)). As will become clear later, a higher Frisch elasticity improves the performance of our model, so in our baseline case we choose the relatively conservative value of 0.3.[18] In the sensitivity analysis, we will experiment with both a higher Frisch elasticity of 0.5 and a case without hours choice (i.e., a 0-1 choice).

Distributions: Learning Ability, Initial Human Capital, and Shocks. Agents have two individual-specific attributes at the time they enter the economy: learning ability and initial human capital endowment. We assume that these two variables are jointly uniformly distributed in the population and are perfectly correlated with each other.[19]Although the assumption of perfect correlation is made partly for simplicity, a strong positive correlation is plausible and can be motivated as follows. The present model is interpreted as applying to human capital accumulation after age 20 and by that age high-ability individuals will have invested more than those with low ability, leading to heterogeneity in human capital stocks at that age, which would then be very highly correlated with learning ability. Indeed,

[18]With our baseline calibration, the Frisch elasticities in the population range from 0.25 to 0.39.

[19]We prefer the uniform distribution over a Gaussian distribution because it has a bounded support, so initial human capital and ability can be easily ensured to be non-negative. Another choice would be a lognormal distribution, but most empirical measures of ability find it more closely approximated by a symmetric distribution, unlike a lognormal one. It will turn out, however, that the *wage* distribution generated by the model will be closer to lognormal with a longer right tail (more consistent with the data), as a result of the convexity arising from the human capital production function.

Table 5: Baseline Parametrization

Parameter	*Description*	*Value*
φ	Curvature of utility of leisure	5.0 (*Frisch = 0.3*)
ψ	Weight on utility of leisure	0.20
α	Curvature of human capital function	0.80
S	Years spent in the labor market	45
T	Retirement duration (years)	20
r	Interest rate	0.02
β	Time discount factor	$1/(1+r)$
χ	Maximum investment time on the job	0.50
δ	Depreciation rate of skills (annual)	1.5%
$E\left[h_0^j\right]$	Average initial human capital (scaling)	4.95
Parameters calibrated to match data targets		
$E\left[A^j\right]$	Average ability	0.190
$\sigma\left(h_0^j\right)/E\left[h_0^j\right]$	Coeff. of variation of initial human capital	0.076
$\sigma\left[A^j\right]/E\left[A^j\right]$	Coeff. of variation of ability	0.408
γ	Dispersion of Markov shock	0.23
p	Transition probability for Markov shock	0.90

Huggett, Ventura, and Yaron (2007) estimate the parameters of the standard Ben-Porath model from individual-level wage data, and find learning ability and human capital at age 20 to be strongly positively correlated (corr: 0.792). Making the slightly stronger assumption of perfect correlation allows us to collapse the two-dimensional heterogeneity in A^j and h_0^j into one, speeding up computation significantly.

Therefore, this jointly uniform distribution of (A^j, h_0^j) yields four parameters to be calibrated. $E\left[h_0^j\right]$ is a scaling parameter and is simply set to a computationally convenient value, leaving three parameters: (i) the cross-sectional standard deviation of initial human capital, $\sigma\left(h_0^j\right)$, (ii) the mean learning ability, $E\left[A^j\right]$, and (iii) the dispersion of ability, $\sigma\left(A^j\right)$. The idiosyncratic shock process, ϵ, is assumed to follow a first-order Markov process, with two possible values, $\{1-\gamma, 1+\gamma\}$, and a symmetric transition matrix: $\Pi = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}$.

This structure yields two more parameters, γ and p, to be calibrated—for a total of five parameters. Finally, because there is measurement error in individual-level wage data, we add a zero mean iid disturbance to the wages generated by the model (which has no effect on individuals' optimal choices).

Data Targets. Our calibration strategy is to require that the wages generated by the model be consistent with micro-econometric evidence on the dynamics of wages found in panel data on US households. Specifically, these empirical studies begin by writing a stochastic process for log wages (or earnings) of the following general form:

$$\log \widetilde{w}_s^j = \underbrace{\left[a^j + b^j s\right]}_{\text{systematic comp.}} + \underbrace{z_s^j + \varepsilon_s^j}_{\text{stochastic comp.}} \tag{16}$$
$$z_s^j = \rho z_{s-1}^j + \eta_s^j,$$

where $\widetilde{w}_s^j$ is the "wage residual" obtained by regressing raw wages on a polynomial in age; the terms in brackets, $[a^j + b^j s]$, capture the *individual-specific* systematic (or life cycle) component of wages that result from differential human capital investments undertaken by individuals with different ability levels, and z_s^j is an AR(1) process with innovation η_s^j. Finally, ε_s^j is an iid shock that could capture classical measurement error that is pervasive in micro data and/or purely transitory movements in wages. For concreteness, in the discussion that follows, we refer to the first two terms in brackets as the "systematic component" and to the latter two terms as the "stochastic component" of wages.

We begin with ε_s and assume that it corresponds to the measurement error in the wage data. This is consistent with the finding in Guvenen and Smith (2009) that the majority of transitory variation in wages is due to measurement error. Based on the results of the validation studies from the US wage data,[20] we take the variance of the measurement error to be 10% of the true cross-sectional variance of wages in each country, which yields $\sigma_\varepsilon^2 = 0.034$ for the United States. We then choose the following five moments from the US data to pin down the five parameters identified earlier:

1. the mean log wage growth over the *life cycle* (informative about $E(A^j)$),
2. the cross-sectional dispersion of wage growth rates, $\sigma(b^j)$ (informative about $\sigma(A^j)$),
3. the cross-sectional variance of the stochastic component (informative about γ),
4. the average of the first three autocorrelation coefficients of the stochastic component of wages (informative about p), and
5. the log 90-10 wage differential in the population (which, together with the previous moments, is informative about $\sigma(h_0^j)$).

[20] For an excellent survey of the available validation studies and other evidence on measurement error in wage and earnings data, see Bound, Brown, and Mathiowetz (2001).

The target value for the mean log wage growth over the life cycle (i.e., the cumulative growth between ages 20 and 55) is 45%. This number is roughly the middle point of the figures found in studies that estimate life cycle wage and income profiles from panel data sets such as the Panel Study of Income Dynamics (PSID); see, for example, Gourinchas and Parker (2002), Davis, Kubler, and Willen (2006), and Guvenen (2007). The second data moment is the cross-sectional standard deviation of wage growth rates, $\sigma(b^j)$. The estimates of this parameter are quite consistent across different papers, regardless of whether one uses wages or earnings (which is not always the case for some other parameters of the income process).[21] We take our empirical target to be 2%, which represents an average of these available estimates.

The next two moments are included to ensure that the model is consistent with some key statistical properties of the *stochastic component* of wages in the data. These moments are (i) the unconditional variance of the stochastic component, $(z_s + \varepsilon_s)$, as well as (ii) the average of its first three autocorrelation coefficients. The empirical counterparts for these moments are taken from Haider (2001), which is the only study that estimates a process for hourly wages *and* allows for heterogeneous profiles. The figure for the unconditional variance can be calculated to be 0.109 and the average of autocorrelations is calculated to be 0.33, using the estimates in Table 1 of his paper.[22,23]

Our fifth, and final, moment is the log 90-10 wage differential in 2003. Adding this moment ensures that the calibrated model is consistent with the overall wage inequality in the US in that year, which is the benchmark against which we measure all other countries. The empirical target value is 1.57 (from the OECD's Labour Force Survey data). Table 6 displays the empirical values of the five moments as well as their counterparts generated by the calibrated model. As can be seen here, all moments are matched fairly well; some

[21]Using male hourly earnings data, Haider (2001) estimates a value of 2.07%, and using annual earnings data he estimates it to be 2.02%. Baker (1997, Table 4, rows 6 and 8) uses an annual earnings measure and estimates values of 1.76% and 1.97% in the two most closely related specifications to the present paper, whereas Guvenen (2009) finds a value of 1.94%, again using male annual earnings data. Finally, Guvenen and Smith (2009) estimate a process for *household* annual earnings and obtain a value of 1.87%.

[22]Over the sample period, Haider estimates the average innovation variance to be 0.074, an AR coefficient of 0.761, and an MA coefficient of -0.42. Using these parameters, the unconditional variance is 0.109.

[23]We match the average of the first three autocorrelation coefficients because Haider (2001) estimates an ARMA(1,1) process, whereas in our model we employ a slightly more parsimonious structure (AR(1)+ iid shock). This latter formulation is a common choice in calibrated macroeconomic models because it requires one fewer state variable while still capturing the dynamics of wages quite well. Nevertheless, because of this difference, it is not possible to exactly match each autocorrelation coefficient in the ARMA(1,1) specification and, so, we match the average of the first three. In the calibrated model, the first three autocorrelations are 0.48, 0.33, and 0.20 compared to 0.42, 0.32, and 0.24 in the data.

Table 6: Empirical Moments Used for Calibrating Model Parameters

Moment	*Data*	*Model*
Mean log wage growth from age 20 to 55	0.45	0.46
Cross-sectional standard deviation of wage growth rates	2.00%	2.00%
Cross-sectional variance of stochastic component	0.109	0.107
Average of first three autocorrelation coeff. of stochastic component	0.33	0.34
Log 90-10 ratio in 2003	1.57	1.58

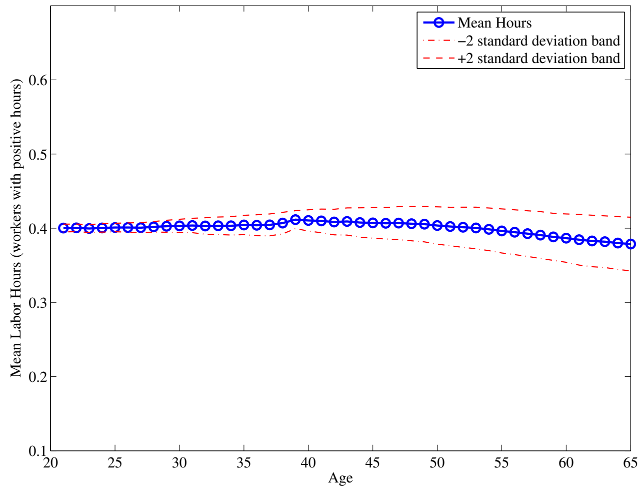
are matched exactly.[24] One point to note is that even though the average of the first three autocorrelation coefficients is pretty low (0.33), recall that the stochastic component includes measurement error as well, which is iid. The Markov shocks to human capital, which approximate the AR(1) process in the data, have a first order annual autocorrelation of 0.80 (implied by $p = 0.90$ shown in Table 5).

Unemployment and Pension System. A great deal of variation can be found across countries in the parameters that control the generosity, the duration, and the insurance component of the benefits system. For example, among the countries in our sample, individuals in Denmark and the Netherlands receive the largest pension payments after retirement with the present value of retirement wealth for the average individual exceeding half a million US dollars (as of 2007). The US and the UK, however, have the lowest pension entitlements—less than six times the average annual earnings in each respective country (and less than half the wealth in Denmark and the Netherlands). We provide the exact formulas for each country and discuss the specifics in more detail in Appendix D. Finally, the calibration of G (the surplus wasted by the government) is challenging because of the difficulty of obtaining reliable estimates of its magnitude. In the baseline case, we assume $G = 0$. So, the government rebates back all the surplus to households in a lump-sum fashion (Tr). We relax this assumption in Section 5.3 and find that it has very little effect on the results.

Consumption Taxes. The average tax rate on consumption is taken from McDaniel (2007), who provides estimates for 15 OECD countries for the period 1950 to 2003 by calculating the total tax revenue raised from different types of consumption expenditures and dividing this number by the total amount of corresponding expenditure. McDaniel

[24]Because the moments chosen are typically non-linear functions of the underlying parameters we calibrate, having five moments and five parameters does not guarantee that all moments will be matched exactly. Considering this, the close correspondence is an encouraging sign that the model is flexible enough to generate wage dynamics similar to that observed in the data.

Figure 5: Life Cycle Profile of Labor Hours



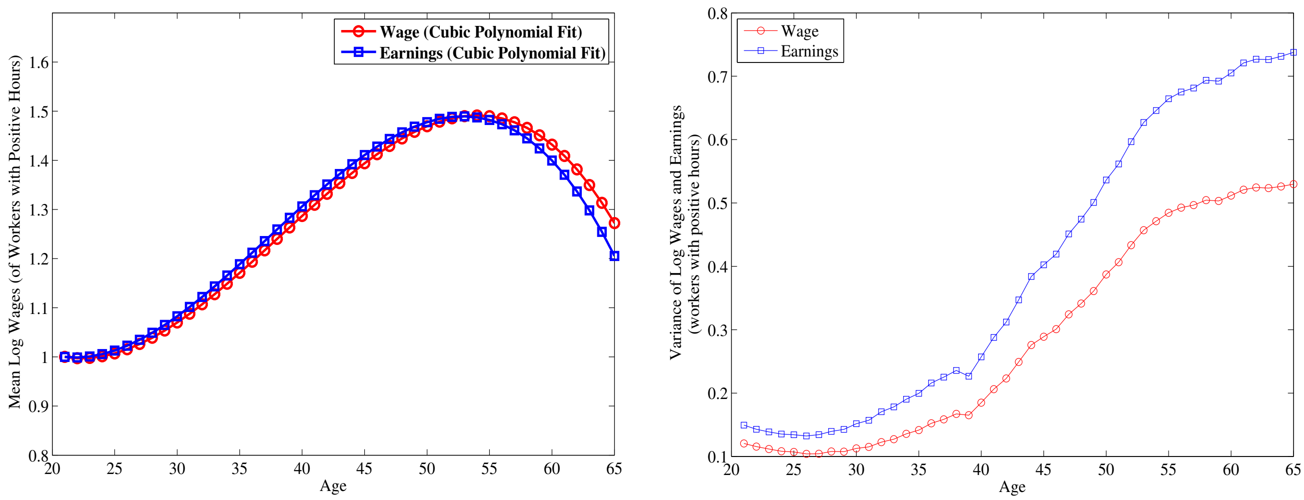
(2007) does not provide an estimate for Denmark, so we set this country's consumption tax equal to that of Finland, which has a comparable value-added tax rate.

4.2 Life Cycle Profiles of Wages, Earnings, and Hours

Before concluding this section, a useful step is to briefly examine if the calibrated model produces plausible behavior over the life cycle for wages, earnings, and labor hours compared to the US data. Figure 5 plots the mean log hours of employed workers, which is computed using 10,000 simulated life cycle paths for individuals drawn from the joint distribution of (A^j, h_0^j). As seen here, the average hours is close to the chosen target of 0.40 and displays little trend over the life cycle. In the US data, average hours rises to about age 25 and then remains fairly flat until about age 55, after which point it starts declining until retirement (cf. Erosa, Fuster, and Kambourov (2009), Figure 2). Although the model does not capture the rise in hours before age 25, the flat hours profile during most of the working life is well captured by the model. Hours also decline in the model, especially after age 50, although not by as much as in the data. Some of the decline in the data is attributable to health shocks or partial early retirement, which are not modeled here.

The dashed lines around the mean profile in Figure 5 show the two standard deviation bands of the hours distribution in the population, which reveal a small rise in hours dispersion over the life cycle. More precisely, the variance of log hours goes up by 1.6 log points, which is fairly small compared to the mean hours of 0.40. Again, this is broadly consistent

Figure 6: Life Cycle Profiles of Wages and Earnings: Mean (Left), Variance (Right)

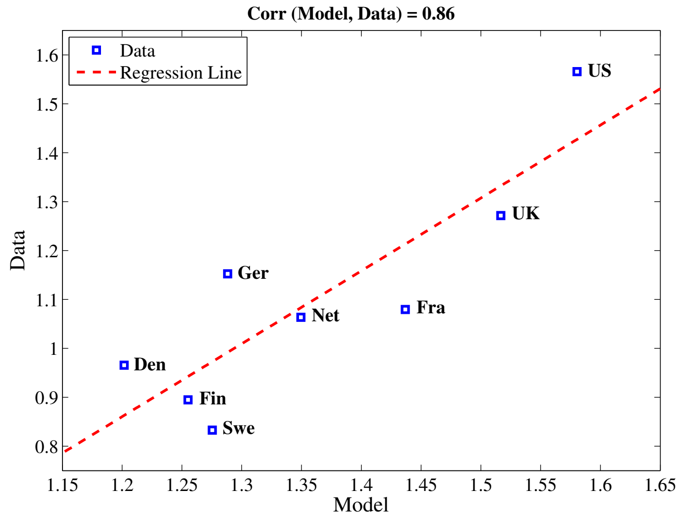


with the findings of Erosa, Fuster, and Kambourov (2009), who document a fairly flat variance profile for hours with a rise only after the mid-40s. Although the rise in the dispersion of hours found by these authors is somewhat larger than what is generated by our model (judging from their Figures 9 and 10), this discrepancy can be fixed here by increasing the Frisch labor supply elasticity, an exercise we conduct in Section 5.3.

The left panel of Figure 6 plots the life cycle profile of average log wages and earnings approximated by a cubic polynomial in age, as commonly done in the literature. The model reproduces the well-known hump shape in wages and earnings. In particular, the mean log wage grows by 45% (as calibrated) and peaks around age 50 and then declines by about 20% until retirement age. Mean log earnings follows a similar pattern but declines by about 5% more than the mean wage as a result of the fall in labor supply later in life, seen in Figure 5. Finally, the right panel plots the variances of log wages and earnings, which both rise in a convex fashion up to 55 and then grow more slowly. The variance of earnings grows faster than that of wages because hours dispersion rises over the life cycle. Guvenen (2009) constructs the empirical counterpart for earnings from the PSID and finds this profile to rise from about 0.20 to 0.73 from age 22 to 62. In the model, the variance rises from about 0.15 to 0.75 from age 20 to 65, fairly consistent with this empirical evidence.

Overall, with the five empirical moments we targeted, the model appears to generate life cycle behavior—in terms of both first and second moments—that is broadly consistent with the data, which is encouraging for the cross-country comparisons we undertake next.

Figure 7: Log 90-10 Wage Dispersion in Eight OECD Countries: Model versus Data



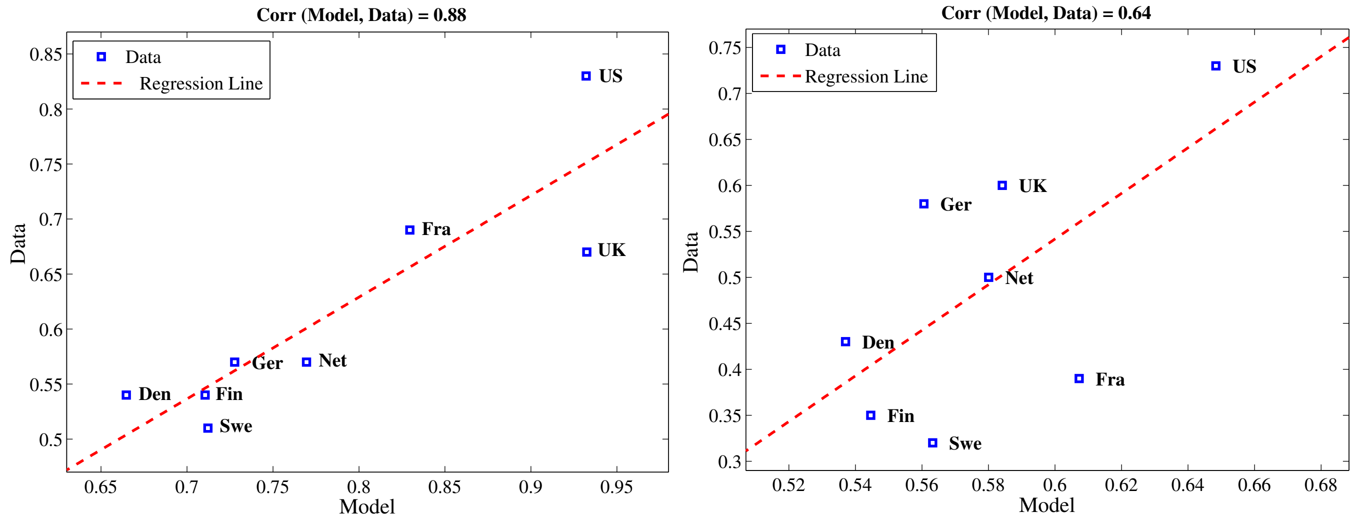
5 Cross-Sectional Results

In this section, we begin by presenting the implications of the calibrated model for wage inequality differences across countries at a point in time. We then provide decompositions that quantify the separate effects of progressivity, average income tax rates, consumption taxes, and benefits institutions on these results. We then perform sensitivity analyses with respect to key parameters.

5.1 The Cross Section in the 2000s

First, Figure 7 plots the log 90-10 wage differential for each country in the data against the value implied by the model. The correlation between the simulated and actual data is 0.86, suggesting that the model is able to capture the *relative ranking* of these eight countries in terms of overall wage inequality observed in the data. Of course, with eight data points a seemingly high correlation can be driven by a few outliers with no obvious pattern among the rest of the data points. As seen in the figure, however, this is not the case: the countries line up nicely along the regression line. Similarly, the left panel of Figure 8 plots the log 90-50 wage differential for each country in the data against the predicted value by the model. The correlation between the actual and simulated data is even higher—0.88—for the log 90-50 wage differential. The correlation of the simulated and actual log 50-10 wage differential, on the other hand, is somewhat lower at 0.63 (right panel of Figure 8). Thus,

Figure 8: Wage Dispersion in Eight OECD Countries: Log 90-50 (left) and Log 50-10 (right) Differentials



the model does a better job in matching the relative ranking of countries for the upper end wage inequality. This finding is consistent with the idea that progressive taxation affects the human capital investment of high-ability individuals more than others and, therefore, the mechanism is more relevant above the median of the wage distribution.

Although these figures and correlations reveal a clear qualitative relationship, they do not allow us to *quantify* how important taxation is for cross-country differences in inequality. For this, we turn to Table 7. The first two columns report the log 90-10 wage differential in the data for all countries, first in levels (second column) and then expressed as a deviation from the US, which is our benchmark country (third column). For example, in Denmark the log 90-10 differential is 0.97, which is 60 log points lower than that in the US. The third and fourth columns display the corresponding statistics implied by the calibrated model. Again, for Denmark, the model generates a log 90-10 differential that is 38 log points below what is implied by the model for the US. Therefore, the model explains 63% (= 38/60) of the difference in the log 90-10 differential between the US and Denmark, reported in column (e). Similar comparisons show that the model does quite well in explaining the level of wage inequality in Germany (41 log points lower than the US inequality in the data versus 29 log points lower in the model) but does poorly in explaining the UK (29 log points difference in the data versus 7 log points in the model). The fraction explained by the model ranges from 30% for France to 71% for Germany. Overall, the model explains 49% of the actual gap in inequality between the US and the CEU in 2003.

Table 7: Measures of Wage Inequality: Benchmark Model versus Data

	Log 90-10					*Log 90-50*	*Log 50-10*
	Data		Model		% explained	% explain.	% explain.
	Level	Δ from US	Level	Δ from US	(d)/(b)		
	(a)	(b)	(c)	(d)	(e)	(f)	(g)
Denmark	0.97	0.60	1.20	0.38	63	93	37
Finland	0.89	0.67	1.26	0.33	49	77	27
France	1.08	0.49	1.44	0.14	30	74	12
Germany	1.15	0.41	1.29	0.29	71	79	59
Netherlands	1.06	0.50	1.35	0.23	46	63	30
Sweden	0.83	0.73	1.28	0.30	42	69	20
CEU	1.00	0.57	1.30	0.28	**49%**	**76%**	**27%**
UK	1.27	0.29	1.51	0.07	22	1	49
US	1.57	0.0	1.58	0.00	–		

To understand which part of the wage distribution is better captured by the model, the next two columns display the same calculation performed in column (e), but now separately for the log 90-50 (f) and 50-10 (g) differentials. For all countries in the CEU, the model explains the upper tail inequality much better than the inequality at the lower end. For example, for Denmark, the model explains 93% of the log 90-50 differential while only generating 37% of the log 50-10 differential. In fact, the model explains *at least* 63% of the upper tail inequality for all countries in the CEU, averaging 76% across all countries, whereas it explains on average only 27% of the log 50-10 differential. Among the CEU, Germany is the one best explained by the model overall: a healthy 79% of the upper tail and 59% of the lower tail inequality is generated by the model. The model does poorly in explaining the small log 50-10 differential in France (12%). One reason could be the legal minimum wage (not modeled here), which is equal to 62% of average earnings in France—the highest among the CEU—and much higher than the 36% of average earnings in the U.S. If these differences were modeled, it could be possible to better reconcile the model with the very small lower tail wage inequality in France. Finally, a notable exception to these generally strong findings is the UK, which is an important outlier: the model explains almost none of the difference between the UK and US at the upper tail (1% to be exact), whereas it explains 49% of the inequality at the lower end. As we shall see later, we found UK to be an outlier along most dimensions this paper attempts to explain and the least well understood economy when viewed through the lens of this model.

Finally, we examine if the calibrated model is broadly consistent with the share of wage inequality accounted for by the upper and lower tails in each region. When the data for all countries in the CEU are aggregated, we find that 57% of the log 90-10 wage dispersion in this region is located above the median and 43% is below the median (statistics not reported in the table to save space). This ratio is very well matched by the model (56.5%), even though no moment from the CEU is used in the calibration. Turning to the US, the upper tail inequality as a fraction of the total is slightly lower than the CEU, at 53%. The model somewhat overstates the inequality at the upper tail (59%) compared to the US data.

5.2 Decomposing the Effects of Different Policies

The baseline model incorporates several differences between the labor market policies of the US and the CEU countries. Here, we quantify the separate roles played by each of these components for the results presented in the previous section. We conduct three decompositions. First, we assume that countries in the CEU have the same benefits institutions as the US but differ in all other dimensions considered in the baseline model. This experiment separates the role of the tax system for wage inequality from that of the benefits system. Second, we also set the consumption taxes of each country equal to that in the US but each country retains its own income tax schedule as in the baseline model. This experiment quantifies the explanatory power of the model that is coming from the income tax system alone. Third, we go one step further and assume that each country keeps the same progressivity of its income tax schedule but is identical in all other ways to the US, including the average income tax rate. This experiment isolates the role of progressivity alone. In each case, we adjust the lump-sum transfers to balance the government's budget.[25]

Table 8 reports the results. First, in column 2, we assume that all countries have the same benefits system as the US. In panel A, the correlation between the data and model is only slightly lower than in the baseline case for the log 90-10 and 90-50 differentials and is, in fact, higher for the 50-10 differential. Turning to panel B, the fraction of the US-CEU difference explained by the model goes down in all cases. For example, for the overall inequality, the explained fraction goes down from 0.49 to 0.44. Therefore, (income and

[25] Adjusting the lump-sum transfers creates an income effect on individuals' choices in addition to the changes in policies considered in each experiment. An alternative would be to keep the lump-sum amount fixed and not balance the budget in each case. We have conducted all three experiments both ways and found only quantitatively minor differences (available upon request), so in the paper we only report the case where the lump-sum amount is adjusted.

Table 8: Decomposing the Effects of Different Policies

	Benchmark	All taxes	Lab. Inc. Tax	Progressivity
Diff. from Benchmark:	(1)	(2)	(3)	(4)
Progressivity	—	—	—	—
Average income taxes	—	—	—	set to US
Consumption tax	—	—	set to US	set to US
Benefits institutions	—	set to US	set to US	set to US
	A. Correlation Between Data and Model			
90-10	0.86	0.82	0.75	0.80
90-50	0.88	0.87	0.83	0.87
50-10	0.63	0.79	0.62	0.60
	B. Fraction of US-CEU Difference Explained by Model			
90-10	0.49	0.44 (90%)[a]	0.37 (76%)	0.33 (68%)
90-50	0.76	0.72 (95%)	0.57 (75%)	0.48 (63%)
50-10	0.27	0.22 (82%)	0.20 (74%)	0.20 (74%)

[a]The numbers in parentheses express the fraction explained by the model in each column as a percentage of the benchmark case reported in column (1).

consumption) taxes together account for 90% (= 44/49) of the model's explanatory power for the overall inequality difference, and the benefits system accounts for the remaining 10%. When we look separately at the tails, we see that the benefits system is less important for inequality at the top (5% of the model's explanatory power) and more important at the bottom (18%). The bottom line is that with tax differences alone, the model generates 72% of the wage inequality differences above the median and 44% half of the difference in overall wage inequality observed in the data between the US and the CEU.

In the next column, we also eliminate the differences in consumption taxes across countries. The model-data correlations go further down but, again, somewhat modestly. In panel B, the explanatory power of the model that is attributable to income taxes alone is roughly 75% for all three measures of wage inequality. The difference between columns 2 and 3 provides a useful measure of the role of consumption taxes: these taxes account for about 14% (= 90% − 76%) of the model's explanatory power for overall wage inequality. Consumption taxes are more important for top-end inequality (20% of the model's total explanatory power) and much less important for the lower-end inequality (8%).

Next, we investigate whether the power of income taxes comes from differences in the average rates across countries or from differences in the progressivity structure. In other words, if continental Europe differed from the US only in the progressivity of its labor income tax system—but had the same *average* tax rate on labor income—how much of the differences in wage inequality found in the baseline model would still remain? To answer this question, we proceed as follows. First, we need to be careful about how we adjust the average tax rate to the US level, because many plausible modifications to the tax structure will simultaneously affect progressivity (as measured, for example, by the wedges). We show in Appendix C.2 how the average income tax rate can be adjusted to any desired rate without affecting progressivity. Then, using these hypothetical tax schedules, we solve each country's problem assuming that all countries have identical labor market policies (set to the US benchmark) and their tax schedules generate the same average tax rate as in the US when using individuals' choices made using the US income tax schedule. In column 4, the correlation between the model and the data changes very little compared to the baseline case reported in column 1, regardless of which part of the wage distribution we look at. In panel B, we see that progressivity alone is responsible for 68% of the explanatory power of the model for the log 90-10 differential. Comparing this to the total effect of taxes (calculated earlier as 75%), it becomes clear that progressivity is the key component of the income tax system that is responsible for understanding wage inequality differences.

In summary, the benefits system and consumption taxes together are responsible for about a quarter of the explanatory power of the model for wage inequality. The more important finding concerns the role of progressivity, which, for all practical purposes, is the key component of the income tax structure for understanding wage inequality differences. Differences in the average income tax rate are the least important among the four types of policy differences we examine in this paper.

5.3 Sensitivity Analysis and Discussion

We now conduct sensitivity analyses with respect to some key parameters of the model. We begin with the Frisch labor supply elasticity and consider two opposite cases: (i) the case with a high Frisch elasticity of 0.5 and (ii) the case without continuous hours choice: $n \in \{0, 0.40\}$. As a third exercise, we allow for the possibility that some of the budget surplus is wasted (i.e., $G > 0$) rather than being rebated back to households in full. In each case that follows, the model is recalibrated to match the same five targets in Table 6. Finally, we discuss the implications of the model for wage inequality among male workers, instead of all workers considered so far.

Table 9: Effect of Labor Supply Elasticity on Wage Inequality Differences

	Frisch = 0.5			*Discrete hours: $n \in \{0, 0.40\}$*		
	Log 90-10	Log 90-50	Log 50-10	Log 90-10	Log 90-50	Log 50-10
	(a)	(b)	(c)	(d)	(e)	(f)
Denmark	84	1.22	48	46	47	47
Finland	62	96	36	34	41	28
France	39	97	16	15	32	9
Germany	85	95	68	45	39	56
Netherlands	54	78	30	28	34	24
Sweden	53	92	24	25	32	20
CEU	**62%**	**97%**	**33%**	**32%**	**38%**	**28%**
UK	32	-10	86	25	0	56

5.3.1 Effect of Labor Supply Elasticity

Frisch Elasticity = 0.5. We begin by setting $\varphi = 3.0$, which implies a Frisch elasticity of 0.5. We then recalibrate the five parameters discussed in Section 4.1 to match the same five moments reported in Table 6. Table 9 reports the counterpart of the analysis we conducted for the benchmark model and reported in Table 7. Comparing the two tables makes it clear that a higher Frisch elasticity improves the model's explanatory power across the board. Now the model can explain 62% of the US-CEU difference in the log 90-10 wage differentials (compared to 49% in the benchmark case) and a remarkable 97% of the upper tail inequality (from 76% before). However, the improvement in the log 50-10 differential is very modest, going up to 33% from 27% in the benchmark case.

Are there any drawbacks to using this alternative calibration? Not many, in fact. One notable shortcoming compared to the benchmark case is that because hours and wages are strongly correlated in this model, a higher Frisch elasticity implies that the rise in earnings inequality is significantly higher than the rise in wage inequality over the life cycle. How important this shortcoming, though, is not obvious. One possible fix would be to add preference shocks to the value of leisure (similar to the route taken by Heathcote, Storesletten, and Violante (2008) but also allow the value of leisure to change stochastically). This approach would reduce the wage-hours correlation and could mitigate this problem without greatly affecting the main results. In this paper, we have not pursued this approach of introducing more heterogeneity into an already rich and complex model and, instead, opted for the more conservative lower Frisch elasticity figures for our baseline model.

Discrete Hours Choice: Full-Time Work versus Unemployment. To better understand the role of continuous labor hours choice, we now examine another case where workers can only choose between full-time employment at fixed hours ($n = 0.40$) and unemployment. The parameters of the utility function are the same as in the baseline case. The results are reported in the last three columns of Table 9. Without the amplification provided by endogenous labor supply—and the resulting dispersion in hours both within each country and across countries—the explanatory power of the model falls and, in some cases, it falls significantly. For example, the model explains 32% of the difference in the log 90-10 differential, compared to 49% in the benchmark case and 62% in the high Frisch case. For the upper-end inequality, the difference is even larger: the model now explains 38%, half of the baseline value, and also much lower than the 97% in the high Frisch case. The difference is much smaller at the lower tail, however, where the explained fraction slightly rises to 28% from 27% in the baseline, and is only a bit lower than 33% in the high Frisch case. Overall, these findings underscore the importance of the interaction of endogenous labor supply choice with progressive taxation for understanding wage inequality differences across countries, especially above the median of the distribution.

5.3.2 Wasteful Government Expenditures versus Transfers

In the baseline model, the surplus was rebated back to households in a lump-sum fashion, essentially assuming that government expenditures are perfect substitutes for private consumption. To examine if our results are sensitive to this assumption, we now assume that half of the government surplus is wasted: $G = Tr$, and each component equals half of the budget surplus (i.e., tax revenues minus benefits payments). This assumption is probably extreme, but it is useful in illustrating whether the results are sensitive to this scenario. From Table 10, we see that, qualitatively, the explanatory power of the model is lower for some countries for the log 90-10 and 90-50 differentials but higher for the 50-10 differential. Quantitatively, however, the effect is minimal across the board. In fact, in some cases, no difference is visible (because of rounding) compared to the benchmark case in Table 7.

5.3.3 Implications for Male Wage Inequality

A potential caveat to the analysis conducted so far is that we study wage inequality among all workers (including females) whereas some of the model parameters were calibrated in Section 4.1 using empirical targets obtained from male wage data. This approach was

Table 10: Effect of Wasteful Government Spending on Wage Inequality Results

	$G = Tr = 0.5\times$ *Gov't Surplus*		
	Log 90-10	Log 90-50	Log 50-10
	(a)	(b)	(c)
Denmark	63	90	38
Finland	49	75	29
France	30	71	14
Germany	69	75	60
Netherlands	45	59	31
Sweden	42	67	23
CEU	**49%**	**73%**	**29%**
UK	21	0	49

necessitated by the fact that reliable empirical estimates of these moments for female workers are difficult to find in the literature (mainly due to the difficulties involved with the extensive margin of female labor supply). To see if this is an important issue, we recalculate the main statistics reported in the previous section, but this time using data from the OECD on *male* wage inequality alone. For male workers, the model explains 45% of the difference between the US and the CEU in the log 90-10 wage differential, 79% of the log 90-50 differential, and 23% of the log 50-10 differential. These figures are quite comparable to those obtained in the baseline model for *all* workers (Table 7). Similarly, the model-data correlations using male data are 0.91 for the log 90-10, 0.86 for the log 90-50, and 0.78 for the log 50-10 wage differentials, which are, again, quite close to those from the baseline model (Figures 7 and 8). Overall, these results suggest that calibrating some parameters by matching moments from male wage data despite our focus on overall wage inequality (because of its obvious relevance for the *overall* state of the economy) is probably not greatly affecting our results. They also to suggest that the human capital channels explored in this paper are likely to be important for female workers as much as for males, given that the model's explanatory power for males and overall wage inequality are very close to each other.

5.4 Other Implications

Besides wage dispersion, the model studied so far also makes predictions for some aggregate variables. We now discuss these briefly (Table 11). Notice that the first two columns report variables only as ratios. This is because for GDP per worker, the levels are not informative (and not comparable to the data counterpart); and for hours per worker, the model was

Table 11: **Aggregate Variables in the CEU and in the US: Model vs Data, 2001–5**

	GDP/ Worker	Hours/Worker	Unemp. Rate		Educational Attainment	
	CEU/US	CEU/US	CEU	US	CEU	US
Data	0.769	0.813	7.4%	5.5%	27.8%	38.0%
Model	0.773	0.905	7.5%	5.7%	21.5%	37.9%

calibrated to match the US data exactly ($n = 0.40$), so, again, there is no information in levels.

The model does a good job of matching GDP per worker differences: in the data, the CEU has a GDP per worker that is 23% lower than that of the US, which is nearly matched by the model. The UK, on the other hand, is again an outlier (not shown in the table). In the model, UK's GDP per worker is only 1.4% lower, whereas it is 24% lower in the data. Turning to hours per worker, the CEU is 19% below the US in the data. The model captures half of this difference and generates 9.5% lower hours per worker for the CEU. Since we allow for unemployment and model the differences in the benefits system, examining the implications of the model along this dimension is also of interest. Somewhat surprisingly, the average unemployment rate in the model is quite close to the data both for the US and the CEU. Again, the UK is an outlier, where the model generates an unemployment rate of 9.3% compared to only 4.8% in the data. Finally, the model also does well in accounting for the educational attainment rate[26] for the US, but underestimates it for the CEU (21.5% compared to 27.8% in the data). Of course, in our model, education is currently treated in a simple manner—as an option for accumulating human capital full time with the same production function used for on-the-job training—so these comparisons should be taken with a grain of salt. A more thorough modeling of the differences in formal education between the US and Europe is a difficult problem, but also a potentially interesting and fruitful direction to extend the current model, which we intend to undertake in future work.

6 Inequality Trends over Time: 1980–2003

In the one-factor model studied so far, the price of human capital, P_H, is simply a scaling factor and has no effect on any implications of the model (which is why we normalized it to

[26]This is defined as the fraction of population aged 24–65 who have completed two or more years of college education (following Autor, Katz, and Kearney (2008)).

1 earlier). In other words, the Ben-Porath framework does not have a well-defined notion of returns to skill. This is an important shortcoming when the goal is to study the changes in human capital behavior over time in response to skill-biased technical change. Guvenen and Kuruscu (2009) proposed a tractable way to extend the Ben-Porath model that allows for a notion of returns to skill and overcomes this difficulty. We now describe the necessary modifications to the model presented earlier.

6.1 Model 2: An Extended Framework

Suppose that individuals now have two factors of production: they begin life with an endowment of "raw labor" (i.e., strength, health, etc.), and, as before, they are able to accumulate human capital over the life cycle. Let l^j denote the initial raw labor of an individual of type j.[27] Raw labor and human capital command separate prices in the labor market, and each individual supplies both of these factors of production at competitively determined wage rates, denoted by P_L and P_H, respectively. Individuals begin their life with zero human capital and each period produce new human capital, Q^j, according to the following generalized Ben-Porath technology:

$$Q^j = A^j \left[(\theta_L l^j + \theta_H h^j) i^j n^j\right]^\alpha . \tag{17}$$

Notice that we now allow both factors of production to affect learning. The motivation for this specification is that an individual's physical capacity (health, strength, stamina, etc.) is also likely to affect her productivity in learning, in addition to her ability and existing human capital stock. Furthermore, Guvenen and Kuruscu (2009) show that this particular specification generates plausible implications for the behavior of wages in the US since the 1970s, which is another reason for adopting this formulation. Finally, both raw labor and human capital depreciate every period at the same rate δ. With this new two-factor structure, the observed *total wage income* of an individual is given by

$$y_s^j \equiv \epsilon \left[P_L l^j + P_H h_s^j\right] n_s^j (1 - i_s^j). \tag{18}$$

Skill-Biased Technical Change. The two-factor structure just introduced breaks the neutrality of the human capital investment with respect to a change in P_H. In particular,

[27] The dependence of raw labor on j makes clear that raw labor can vary across individuals, albeit in a way that is perfectly correlated with ability. We provide justification for this structure later on.

now a rise in P_H increases investment, even when P_L is fixed. Before delving into the quantitative results, it is useful to step back and understand this point more clearly. To this end, we make several assumptions that yield an analytical expression for the optimality condition.[28] In addition, we also assume $P_H/P_L = \theta_H/\theta_L$, which essentially means that the relative price of human capital to raw labor is the same as their relative productivity in the human capital function. Although this assumption is not necessary for quantitative results, following Guvenen and Kuruscu (2009) we make it in the rest of the paper because it substantially simplifies the solution of the model. Under these assumptions, the first-order condition is

$$C_j'(Q_s^j) = \theta_H \{\beta \frac{1-\tau(y_{s+1})}{1-\tau(y_s)} n_{s+1} + \beta^2 \frac{1-\tau(y_{s+2})}{1-\tau(y_s)} n_{s+2} + ... + \beta^{S-s} \frac{1-\tau(y_S)}{1-\tau(y_s)} n_S\} \tag{19}$$

The key observation is that optimal investment, Q_s^j, now depends on the level of θ_H, unlike in (4) and (6) presented in Section 2.1, where P_H did not appear at all. This is because, in our two-factor model, the marginal cost of investment depends on both θ_H and θ_L (see equation (18)), whereas the marginal benefit is only proportional to θ_H. As a result, a higher θ_H (for example, due to SBTC) increases the benefit more than the cost (since θ_L does not rise), resulting in higher investment. This feature is an important difference between this two-factor model and the standard Ben-Porath framework.

6.2 Results: US versus CEU with *Fixed* Tax Schedules

The extended model has some new parameters that need to be calibrated. Except those discussed here, all parameter values are kept at the values given in Table 5. An important point to note is that for the cross-sectional analysis of the previous section, the two-factor model would have precisely the same implications as the one-factor Ben-Porath model used earlier. This is because θ_H and θ_L are constant at a point in time and their values can be normalized to generate exactly the same results as in the previous section. Thus, with proper choices of θ_H, θ_L, and the distribution of l^j, we do not need to recalibrate any other parameter and can still obtain the same results for year 2003 as before. This is the route that we follow in this section.[29]

[28] We set $\chi \equiv 1$, eliminate the benefits system ($\Omega \equiv 0$ and $\Phi \equiv 0$), and set $\epsilon = 1$. For simplicity (although not necessary), we also set $\delta = 0$.

[29] More specifically, the two-factor model eliminates initial heterogeneity in human capital but instead introduces raw labor. We make the same assumptions for l^j as we made earlier about h_0^j. That is, we assume

Table 12: **Rise in Wage Inequality: Model versus Data, 1980–2003.** The model is calibrated to match the 23 log points rise in the log 90-10 differential for the US from 1980 to 2003.

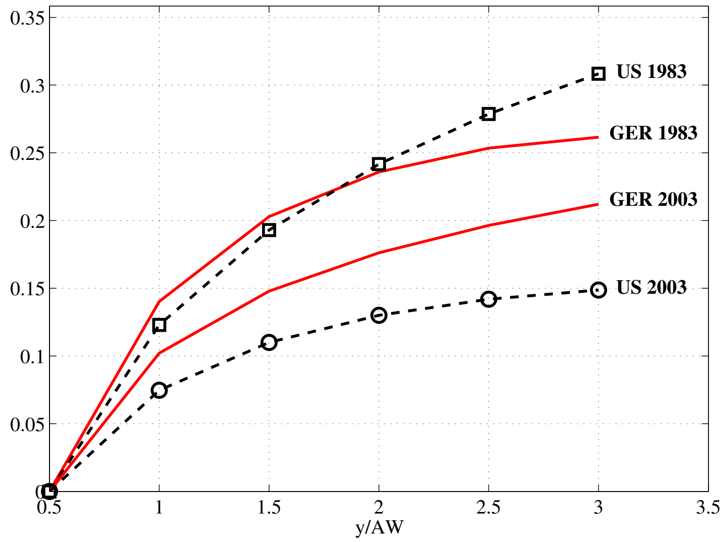
			Change in Log Wage Differentials				
			Log 90-10	=	*Log 90-50*	+	*Log 50-10*
CEU	Data	*Level*	0.070		0.063		0.007
		%			91%		9%
	Model	*Level*	0.168		0.129		0.039
		%			77%		23%
US	Data	*Level*	0.230		0.160		0.070
		%			70%		30%
	Model	*Level*	0.232		0.184		0.048
		%			79%		21%
Difference	Data:	*Level*	0.168		0.097		0.063
		%			61%		39%
	Model	*Level*	0.065		0.056		0.009
		%			87%		13%
% Explained			**41%**		**58%**		**14%**

For examining the change in inequality over time, we choose $\Delta \log (\theta_H/\theta_L)$ to match the 23 log points increase in the log 90-10 wage differential in the US from 1980 to 2003. The required change in $\Delta \log (\theta_H/\theta_L)$ is 0.236. With this calibration, wage inequality rises by 0.168 in CEU during the same time, compared to 0.070 rise in the data (fourth column of Table 12). These results imply that differences in labor market policies, even when they are fixed over time, can generate about 41% ($= (0.232-0.168)/(0.230-0.070)$) of the widening in the inequality gap between the US and the CEU during this time period.

Another dimension of the rise in wage inequality is seen in Table 2 and replicated in the last two columns of Table 12. The substantial part of the rise in wage inequality in the CEU has been at the top: the log 90-50 differential is responsible for 91% of the total rise in the 90-10 differential, whereas only 9% of the rise took place at the lower end. A similar outcome, somewhat less extreme, is observed in the US where 70% of the rise in the log 90-10 differential is due to the 90-50 differential. The model generates a similar picture: about 77% of the rise in the CEU and 79% in the US is due to the 90-50 differential. An alternative way to express these figures is that the model explains 58% of the increase in the inequality gap above the median between the US and the CEU but only 14% of the rising gap below the median. As is clear by now, this is a recurring theme in this paper:

that l^j is uniformly distributed and is perfectly correlated with A^j. We also assume that $\theta_H = \theta_L = 1$ in 2003, which allows us to use the same mean value and coefficient of variation for l^j as for h_0^j in Table 1.

Figure 9: Progressivity Wedges at Different Income Levels: US vs. Germany, 1983 and 2003



the model explains cross-country inequality facts at the upper tail quite well, but explains a smaller fraction at the lower tail.

6.3 Results: US versus Germany with *Changing* Tax Schedules

For the United States and Germany, we were able to construct the effective tax schedules for 1983, which allows us to conduct a two-country comparison in the presence of both SBTC *and* changing tax schedules. The procedure for constructing the 1983 tax schedules is described in Appendix C.3 and the resulting progressivity wedges are shown in Figure 9. As seen in the figure, in 1983 the progressivity of the tax structure in the US and Germany was similar in both countries up to about twice the average earnings level. And above this point, the US actually had the more progressive system. Over time, the US has become much less progressive, whereas the change in Germany has been more gradual, making the US tax schedules much flatter than that of Germany over time.

Using these schedules, we conduct two experiments.[30] First, we consider the case where there is no SBTC between 1983 and 2003 and the only change has been in the tax schedules (including the consumption tax rates). No parameter is recalibrated to match any target

[30]Because of computational burden, these experiments only provide steady state comparisons. Although solving for the full transition path is beyond the scope of this paper, it could be important for the quantitative results, so future work on this issue is certainly warranted.

Table 13: US vs Germany: Log 90-10 Differential with Changing Tax Schedules

	Data	Model		
		Baseline	Experiment 1	Experiment 2
Taxes		Fixed	*Changing*	*Changing*
θ_H (SBTC)		Calibrated to US	Fixed	Calibrated to US
US	0.23	0.232	0.167	0.230
Germany	0.09	0.154	0.010	0.068
% Explained		55%	—	116%

in 1983. The results are reported in the fourth column of Table 13 (denoted Experiment 1). In the US, the log 90-10 differential rises by 16.7 log points compared to 23 log points in the data. Hence, the flattening of the tax schedule alone explains a significant fraction (about 72%) of the rise in the US wage inequality during this time. To our knowledge, this result is new to this paper. In contrast to the US, wage inequality barely changes (by +1 log point) in Germany from 1983 to 2003. Thus, the dramatic fall in progressivity in the US and the small change in Germany alone could explain an important part of the difference in the evolutions of inequality between these two countries.

As a second experiment, in addition to changing tax schedules, we now also calibrate the change in the skill bias of technology such that we exactly match the log 90-10 wage differential in the US in 1983. The required change in $\log(\theta_H/\theta_L)$ is 7.5 log points, which is about a third of the value in the baseline model (23.6 log points). Since the model is calibrated to exactly match the US wage inequality, we turn to Germany: the log 90-10 differential rises by less than 7 log points compared to the 9 log points rise in the data. Thus, the model easily generates—in fact, it over-explains by 16% (i.e., $(0.23 - 0.068)/(0.23-0.09) = 1.16$)—the growth of the inequality gap between the US and Germany. For comparison, the baseline model (third column of Table 13) with fixed tax schedules explained about 55% of the rise in the inequality gap between the US and Germany.

Although these results are certainly encouraging, a caveat must be noted. First, wage inequality in 1983 depends not only on the tax schedule in 1983 but also on those that were in place several years prior, since the dispersion in human capital across individuals results from investments made in previous years. Clearly, the same comment applies to 2003. Although in our exercise we do not account for this fact, it is not clear which way this biases the results. This is because the US tax system was even more progressive before the Economic Recovery Tax Act of 1981, whereas the progressivity change in the years preceding 2003 (say, from 1990 to 2003) was more modest. Therefore, if we were to use a time average of tax schedules in our exercise (say, 1973 to 1983 and 1993 to 2003), we conjecture that the reduction in progressivity over time could be larger than we assumed in the experiment just described (which would attribute an even larger role to taxes).

7 Conclusions

In this paper, we have studied the effects of progressive labor income taxation on wage inequality when a major source of wage dispersion is differential rates of human capital accumulation. To understand the main mechanisms and their quantitative importance, we have examined the inequality differences between the US and the CEU, which differ significantly in their income tax structures as well as in other dimensions of their labor market institutions. A common theme that permeates all of our findings is that the model is significantly better at explaining inequality differences at the upper tail compared to the lower tail. Institutions, such as unionization, minimum wage laws (as in the case of France, discussed earlier), and centralized bargaining, are likely to be more important for the lower tail. However, since changes in the upper tail have been so important during this time (as we have documented), the mechanisms studied in this paper provide a promising direction for understanding US-CEU differences in wage inequality.

We also found that the most important policy difference for wage inequality is the progressivity of the income tax system, which is responsible for about two-thirds of the model's explanatory power. In addition, endogenous labor supply plays an important amplification role for wage inequality when interacted with progressivity.

There is an active debate in the literature on the appropriate values of the labor supply elasticity at the macro and micro levels. This paper also has implications for this issue. In particular, if a micro-econometrician were to use the standard empirical regression of labor hours on wages to recover the Frisch elasticity (of labor supply) using simulated data from our model, she would obtain a value of 0.14 in the baseline case when the theoretical value we set is 0.30 (and 0.26 when the theoretical value is set to 0.50). At the same time, the labor supply elasticity relevant for human capital accumulation is higher than the theoretical value because of the amplification channel discussed earlier. We have also investigated the welfare differences across countries resulting from labor market institutions. The calibrated model implies that for highest-ability individuals, living in the US yields the highest utility, whereas for the lowest-ability individuals, the UK is the best country. For brevity's sake, we leave out of the paper a more detailed discussion of these welfare experiments as well as the micro elasticity regressions (results available upon request).

Finally, we have also investigated if the differential rise in wage inequality between these two regions could be explained by the channels explored here. Using fixed tax schedules over time, the model explains about 40% of the rise in the inequality gap between the US and the CEU and 60% of the upper tail inequality. In a two-country comparison, we found that the model explains all of the rise in the inequality gap between the US and Germany, when the actual changes in the tax schedules were also incorporated.

The model studied in this paper also has implications for how the tax and benefits systems affect the *life cycle* profile of wages and it would be illuminating to examine how these implications compare to the data from the US and the CEU. Such an analysis, however, requires *panel data* on wages (to disentangle age effects from time or cohort effects), which is difficult to obtain on a comparable basis for more than a few countries. One piece of evidence is available from a Swedish panel dataset, studied by Domeij and Floden (2009). These authors find that the dispersion of wage earnings growth rates over the life cycle ($\sigma(b^j)$ defined in Section 4.1) is much smaller in Sweden than in the United States. Given the high progressivity of income taxes in Sweden compared to the US, this outcome is exactly what is predicted by the present model. An interesting research avenue would be to conduct a fuller investigation of the life cycle pattern of wages (and also perhaps hours) in a broader set of countries.

In this paper, we made several assumptions to make the quantitative exercise computationally feasible.[31] As noted earlier, an important direction to extend the current framework would be by carefully modeling the differences between the US and the CEU in the financing of their education systems as well as in the kinds of skills taught in schools in both places. This is a difficult but interesting question that is at the top of our future research agenda.

[31] The numerical solution of the model requires care because the individuals' dynamic problem has several sources of non-convexities. As a result, solving for the equilibrium takes about 14 hours for the US and UK, and as much as 30 hours for some countries like Denmark. This makes calibration very time consuming, which prevented us from extending the model in other directions.

ADDITIONAL MATERIAL

Table A.1: Correlation between Different Labor Market Institutions

	Union density	Union coverage	Centralization & Coordination	PW
Union coverage	0.49	1		
C&C	0.57	0.75	1	
PW	**0.88**	**0.75**	**0.78**	1
PW*	**0.81**	**0.85**	**0.69**	0.93

A Appendix: Progressivity versus Other Labor Market Institutions

Table A.1 reports the cross-correlations between different labor market institutions in a country and the progressivity of its income tax structure. The progressivity measures we use are PW and PW* defined in the text. The labor market institutions are union density, union coverage rate, and C&C (Centralization & Coordination) score. All three definitions are explained in more detail later. All three variables are measured in a way that higher numbers indicate more deviation from a frictionless economy. The main finding is that both measures of progressivity are strongly positively correlated with all three labor market institutions. Therefore, countries that have a more unionized labor force with stronger centralized bargaining are also those that have a more progressive labor income tax system. To our knowledge, this finding is new to this paper.

Definition of Labor Market Institutions.

Union density is commonly measured by the percentage of salaried workers who are union members. The results of collective bargaining agreements between unions and employers are often extended (through mandatory and/or voluntary mechanisms) to non-union workers and firms. The total fraction of workers covered through such extensions is termed *union coverage*. *Centralization* is a measure that indicates the level at which negotiations take place, such as at firm or plant level (i.e., decentralized bargaining), industry level, and countrywide level (centralized bargaining). In many countries, informal networks and intensive contacts between social partners coordinate the behavior of trade unions and employers' associations. Examples are the leading role of a limited number of key wage settlements in Germany, and the active role of powerful employer networks in Japan. Therefore, what matters is not only the formal degree of centralization, but also the degree of informal consensus seeking between bargaining partners. This is generally called the level of *coordination*. The C&C score is an index that increases with the level of centralization and coordination. (Definitions summarized from Borghijs, Ederveen, and de Mooij (2003).)

B Key Derivations and Definitions

B.1 Derivation of the Optimal Investment Condition (eq. (19))

Here, we derive the optimal investment condition in the most general framework studied in this paper (equation (19) in Model 2). The optimality conditions presented earlier in the paper ((3), (4), and (6)) can all be obtained as special cases of this formulation.

Under the assumptions stated in Section 6 (i.e., setting $\chi \equiv 1$, eliminating unemployment benefits and pension payments ($\Omega \equiv 0$ and $\Phi \equiv 0$), and setting idiosyncratic shocks to their mean value), the problem of the agent is given by

$$\begin{aligned} V(h,a,s) &= \max_{c_s,n_s,Q_s} u((1+r)a_s + y_s(1-\bar{\tau}(y_s)) - a_{s+1}, 1-n) \\ &+ V(h_{s+1}, a_{s+1}, s+1) \end{aligned}$$

$$\text{s.t.} \qquad y_s = (\theta_L l + \theta_H h_s) n_s - C(Q_s).$$

Note that total tax liability of the agent is given by $y\bar{\tau}(y)$. The derivative of tax liability with respect to y gives the marginal tax rate. Thus, $\tau(y) = \bar{\tau}(y) + y\bar{\tau}'(y)$. Using this expression, we obtain the following FOCs for this problem

$$\begin{array}{rc} (n_s): & (\theta_L l + \theta_H h_s)(1-\tau(y_s)) u_1(c_s, 1-n_s) = u_2(c_s, 1-n_s) \\ (a_s): & u_1(c_s, 1-n_s) = \beta V_2(h_{s+1}, a_{s+1}, s+1) \\ (Q_s): & C'(Q_S)(1-\tau(y_s)) u_1(c_s, 1-n_s) = \beta V_1(h_{s+1}, a_{s+1}, s+1) \end{array}$$

Envelope conditions are:

$$\begin{array}{rc} (a_s): & V_2(h_s, a_s, s) = (1+r) u_1(c_s, 1-n_s) \\ (h_s): & V_1(h_s, a_s, s) = n_s (1-\tau(y_s)) u_1(c_s, 1-n_s) + n_{s+1}\beta V_1(h_{s+1}, a_{s+1}, s+1). \end{array}$$

Combining the envelope conditions with the FOCs yields

$$\begin{aligned} C'(Q_s)(1-\tau(y_s)) = \theta_H n_{s+1}(1-\tau(y_{s+1})) & \underbrace{\frac{\beta u_1(c_{s+1}, 1-n_{s+1})}{u_1(c_s, 1-n_s)}}_{\frac{1}{1+r}} + \\ \theta_H n_{s+1}(1-\tau(y_{s+1})) & \underbrace{\frac{\beta^2 u_1(c_{s+2}, 1-n_{s+2})}{u_1(c_s, 1-n_s)}}_{\frac{1}{(1+r)^2}} + ... \end{aligned}$$

Rearranging this expression delivers equation (19):

$$C_j'(Q_s^j) = \theta_H \{ \beta \frac{1-\tau(y_{s+1})}{1-\tau(y_s)} n_{s+1} + \beta^2 \frac{1-\tau(y_{s+2})}{1-\tau(y_s)} n_{s+2} + ... + \beta^{S-s} \frac{1-\tau(y_S)}{1-\tau(y_s)} n_S \}.$$

B.2 Definition of y^* Introduced in Section 3.1

Recall that y^* was defined in Section 3.1 as "the income an individual would have earned in a economy identical to the present model, except that unemployment insurance was set to zero. Mathematically, the definition is $y^* = h(1-i^*)n^*$, where n^* and i^* are given by the solution to the following problem:

$$
\begin{aligned}
(c^*, n^*, i^*, a'^*(\epsilon')) &= \arg \max_{c,n,i,a'(\epsilon')} \left[u(c,n) + \beta \sum_{\epsilon'} \Pi(\epsilon' \mid \epsilon) V(\epsilon', a'(\epsilon'), h', m+1; s+1) \right] \\
\text{s.t.} \quad (1+\bar{\tau}_c)c + \sum_{\epsilon'} q(\epsilon' \mid \epsilon) a'(\epsilon') &= (1 - \bar{\tau}_n(y))y + a + Tr \\
y &= [\epsilon h(1-i)]n \\
h' &= (1-\delta)h + A(hin)^\alpha, \\
i &\in [0, \chi].
\end{aligned}
$$

C Country-Specific Tax Schedules

C.1 Estimating Country-Specific Average Tax Schedules

Here we provide more details on the estimation of tax schedules described in Section 2.2. Define normalized income as $\widetilde{y} \equiv y/AW$. For each country, denote the top marginal tax rate with τ_{TOP} and the top bracket $\widetilde{y}_{TOP}$. The values for these variables are taken from the OECD tax database.[32] As noted in the text, we already have average tax rates for all income levels below 2 (i.e., two times AW). For values above this number, we have to consider separately the case where a country's top marginal tax rate bracket is lower and higher than 2. In the former case ($\widetilde{y}_{TOP} < 2$), since we know the average tax rate at $\widetilde{y} = 2$, each additional dollar up to 2 is taxed at the rate of τ_{TOP}. Therefore, for $\widetilde{y} > 2$

$$
\bar{\tau}(\widetilde{y}) = (\bar{\tau}(2) \times 2 + \tau_{TOP} \times (\widetilde{y} - 2))/(\widetilde{y})
$$

If instead $\widetilde{y}_{TOP} > 2$ (which is only the case for the US and France), we do not know the marginal tax rate between $\widetilde{y} = 2$ and $\widetilde{y}_{TOP}$. Thus, we first set $\tau(2) = (\bar{\tau}(2) \times 2 - \bar{\tau}(1.75) \times 1.75)/0.25$ and use linear interpolation between $\tau(2)$ and τ_{TOP}. We have

$$
\tau(\widetilde{y}) = \begin{cases} \tau(2) + \frac{\tau_{TOP} - \tau(2)}{y_{TOP} - 2}(\widetilde{y} - 2) & \text{if } 2 < \widetilde{y} < \widetilde{y}_{TOP} \\ \tau_{TOP} & \text{if } \quad \widetilde{y} > \widetilde{y}_{TOP}. \end{cases}
$$

Then the average tax rate function for $\widetilde{y} > 2$ is

$$
\bar{\tau}(\widetilde{y}) = \begin{cases} (\bar{\tau}(2) \times 2 + \tau(\widetilde{y}) \times (\widetilde{y} - 2))/\widetilde{y} & \text{if} \quad 2 < \widetilde{y} < \widetilde{y}_{TOP} \\ (\bar{\tau}(2) \times 2 + \frac{(\tau(2) + \tau_{TOP})}{2}(\widetilde{y}_{TOP} - 2) + \tau_{TOP} \times (\widetilde{y} - \widetilde{y}_{TOP}))/\widetilde{y} & \text{if} \quad \widetilde{y} > \widetilde{y}_{TOP} \end{cases}
$$

We use this expression to compute $\bar{\tau}$ for $\widetilde{y} = 3, 4, ..., 8$ (in addition to the original average tax rate from OECD website). We then fit the functional form given in equation (8) to these 13 data points as explained in the text. The resulting coefficients are reported in Table A.2.

[32] From Table I.7, available for download at www.oecd.org/ctp/taxdatabase.

Table A.2: Tax Function Parameter Estimates

	$\bar{\tau}(y/AW) = a_0 + a_1(y/AW) + a_2(y/AW)^{\phi}$				
Country:	a_0	a_1	a_2	ϕ	R^2
Denmark	1.4647	$-.01747$	-1.0107	$-.15671$	0.990
Finland	1.7837	$-.01199$	-1.4518	$-.11063$	0.999
France	0.5224	.00339	$-.24249$	$-.41551$	0.993
Germany	1.8018	$-.01708$	-1.3486	$-.11833$	0.992
Netherlands	3.1592	$-.00790$	-2.8274	$-.03985$	0.984
Sweden	9.1211	$-.00762$	-8.7763	$-.01392$	0.985
UK	0.5920	$-.00390$	$-.32741$	$-.30907$	0.989
US	1.2088	$-.00942$	$-.94261$	$-.10259$	0.993

C.2 Deriving Tax Schedules with Different Progressivity but Same Average Tax Rate

To change the average tax rates in Europe without changing progressivity, we apply the following procedure. Let $\tau_i(y)$ be the marginal tax rate in country i for income level y. We would like to obtain a new tax schedule $\tau_i^*(y)$ with the same progressivity but with a different level. Thus, we need to have (for all y and y')

$$\frac{1-\tau_i^*(y')}{1-\tau_i^*(y)} = \frac{1-\tau_i(y')}{1-\tau_i(y)} \;\Rightarrow\; \frac{1-\tau_i^*(y')}{1-\tau_i(y')} = \frac{1-\tau_i^*(y)}{1-\tau_i(y)}$$

Letting this ratio to be equal to a constant k, the new tax schedule τ^* is obtained by the following expression:

$$1-\tau_i^*(y) = k(1-\tau_i(y)) \text{ for all } y. \tag{20}$$

Let the average tax rate be

$$\bar{\tau}_i(y) = a_0 + a_1 y + a_2 y^{\phi} \quad \Rightarrow \quad \tau_i(y) = a_0 + 2a_1 y + a_2(\phi+1)y^{\phi}.$$

Plugging this last expression into (20) and solving for $\tau^*(y)$, we get

$$\tau_i^*(y) = 1 - k + k\left[a_0 + 2a_1 y + a_2(\phi+1)y^{\phi}\right].$$

Observing that $y\bar{\tau}_i(y) = \int_0^y \tau_i(x)dx$, we can solve for the average tax rate $\bar{\tau}_i^*(y)$ as

$$\bar{\tau_i}^*(y) = 1 - k + k[a_0 + a_1 y + a_2 y^{\phi}] = 1 - k + k\bar{\tau}_i(y). \tag{21}$$

The new schedule $\bar{\tau}_i^*(y)$ has the same progressivity as $\bar{\tau}_i(y)$ but can have any desired average tax rate. We choose k so that the average labor income tax rate in country i is equal to the average labor income tax rate in the US.

C.3 Constructing Tax Schedules for 1983

Here, we describe the formulas we use to calculate the average tax rate at different income levels for Germany and the United States in 1983. This information is obtained from the OECD (1986) (see pages 104–105 and 244–248 for the US and pages 74–75 and 149–154 for Germany. In all calculations for Germany, the monetary figures are in Deutsche Mark (DM). Gross income is denoted by `GM`.

C.3.1 Germany

Social Security Contributions. In 1983, the social security system in Germany had two brackets with their respective tax rates. Specifically, social security contributions (SSC) were given by:

$$SSC = 0.1138 \times (\min(\texttt{GI}, 64800) + 0.0588(\min(\texttt{GI}, 48600)).$$

Allowances. Each worker receives an allowance (tax exemption) of DM 1080 and an allowance of DM 564 for work-related expenses. The OECD considers other miscellaneous allowances in the amount of DM 1606. We treat this amount as fixed for all levels of income. Finally, workers are able to deduct part of their social security contributions determined by this formula:

$$\begin{aligned}\texttt{SSC Allowance} \quad = \quad & \max\{6000 - 0.18(\texttt{GI}), 0\} \\ & + \min(2340, \max\{SSC - \max\{6000 - 0.18(\texttt{GI}), 0\}\}) \\ & +0.5 \times \min(2340, \max\{SSC - \max\{6000 - 0.18\texttt{GI}, 0\} - 2340, 0\}).\end{aligned}$$

Total Tax. Putting together the taxes and allowances just described gives the taxable income of a worker:

$\texttt{Taxable Income} = \texttt{GI-SSC Allow.-Basic Allow.-Work-related and other Allow.}$

Now, we can calculate the tax liability to the household. The first step is to round the taxable income.

$\texttt{Rounded Taxable Income (RTI)} = round(\texttt{Taxable Income}/54) \times 54.$

We calculate two variables Y and Z that will be used in the calculations that follow. They are defined as $Y = \frac{\texttt{RTI}-18000}{10000}$ and $Z = \frac{\texttt{RTI}-60000}{10000}$. To obtain the income tax for a worker, we need to apply Germany's tax schedule in 1983:

$$\texttt{Income Tax=}\begin{cases} \texttt{zero} & \text{if } \texttt{RTI} \leq 4212 \\ 0.22 \times \texttt{RTI} - 926 & \text{if } 4213 < \texttt{RTI} \leq 18035 \\ (((3.05Y - 73.76)Y + 695)Y + 2200) \times Y + 3034 & \text{if } 18036 < \texttt{RTI} \leq 60047 \\ (((0.09Z - 5.45)Z + 88.13)Z + 5040) \times Z + 20018 & \text{if } 60048 < \texttt{RTI} \leq 130031) \\ 0.56 \times \texttt{RTI} - 14837 & \text{if } \texttt{RTI} > 130032 \end{cases}$$

$$\texttt{Average Tax Rate} = \frac{\texttt{Income Tax} + SSC}{\texttt{Gross Income}}.$$

C.3.2 The United States

Social Security Contribution. In 1983, the employee social security contribution in the US was given by

$\texttt{SSC Employee} = 0.067 \times (\min(\texttt{Gross Income}, 35700))$

The employer's social security contribution matches the employee's contribution of 6.7% on earnings up to \$35700. Additionally, employers are required to pay an unemployment tax of 6.2% of earnings up to \$7000 and a nationwide average for state-sponsored tax plan of 2.8% of earnings up to \$7624.

$$\texttt{SSC Employee} = 0.067 \times (\min(\texttt{GI}, 35700)) + 0.062 \times (\min(\texttt{GI}, 7000)) + 0.028 \times (\min(\texttt{GI}, 7624))$$

Allowances. The total combined allowances and exemptions amount to \$2300 per worker.

$\texttt{Taxable Income} = \texttt{Gross Income} - \texttt{Basic Allowance} - \texttt{Tax Bracket Allowance}.$

Federal Income Tax. Now, we can calculate the tax liability for the household. We need to apply the US tax schedule in 1983. The first \$2300 is not taxed, as discussed earlier. The tax rate is 11% when taxable income is in range $(2300, 3400)$; is 13% in range $(3400, 4400)$; is 15% in range (4400,8500); 17% in range $(8500, 10800)$; is 19% in range (10800,12900); is 21% in range $(12900, 15000)$; is 24% in range (15000,18200); is 28% in range $(18200, 23500)$; is 32% in range (23500,28800); is 36% in range (28800,34100); is 40% in range (34100,41500); is 45% in range (41500,55300); and 50% above \$55,300.

State and Local Taxes. For the purposes of calculating local and state taxes, the OECD considers a worker that lives in Detroit, Michigan. Detroit allows an exemption of \$600, then a flat 3% tax is applied. $\texttt{Tax Detroit} = 0.03(\texttt{GI} - 600)$. The formula for Michigan's state income tax is given by

$\texttt{Tax Michigan} = 0.0635(\texttt{GI} - 1500) - 0.05 \max(\texttt{Tax Detroit-200}, 0) + 27.5$

$\texttt{Total Local Tax} = \texttt{Tax Michigan} + \texttt{Tax Detroit}$

Total Tax. The total tax liability is equal to the income tax plus the social security contribution and the local tax. Then, we have

$$\texttt{Average Tax Rate} = \frac{\texttt{Total Tax Liability}}{\texttt{Gross Income}}$$

D Pension and Unemployment Benefits Systems

Pension System. The details of the pension benefits system for OECD countries used in this paper are taken from the OECD publication entitled "Pensions at a Glance: 2007." The specific numbers used in this section are from Table I.2 and the unnumbered table on page 35 of that document. Further details of these pension systems, including the number of years required to qualify for full benefits, and so on, are described more fully on pages 26–35 of the same document. Let $\overline{y}^j$ be the lifetime average of *net* (after-tax) labor earnings of all individuals with ability level j; and let $\overline{y}$ be the same variable averaged across all ability levels. Finally, recall that m^R is the total

Table A.3: Pension System Formulas

	a	b	Ranges	Ceiling for Pensionable Income (as % of AW)
DEN	0.371	0.528	all	—
FIN	0.011	0.695	all	—
FRA	0.141	0.484	all	300%
GER	-0.004	0.621	if $\overline{y}^j \leq 1.5\bar{y}$	
	0.927		if $\overline{y}^j > 1.5\bar{y}$	150%
NET	0.005	0.928	all	—
SWE	-0.021	0.735	all	367%
UK	0.257	0.154	if $\overline{y}^j \leq \bar{y}$	115%
	0.315	0.096	if $\bar{y} < \overline{y}^j \leq 1.5\bar{y}$	
	0.396	0.042	$\overline{y}^j > 1.5\bar{y}$	
US	0.168	0.355	all	290%

number of years a worker has been employed up to the retirement age, and let $\overline{m}$ be the maximum number of years of work that an individual can accumulate retirement credits in a given country. The *net* retirement earnings of individual with ability j is given as

$$\Omega(\overline{y}^j, m^R) = min\left(1, \frac{m^R}{\overline{m}}\right)\left[a\overline{y} + b\overline{y}^j\right]$$

The first term approximates the credit accumulation process whereby individuals qualify for full retirement benefits after working a certain number of years and only qualify for partial pensions if they retire before that. We set $\overline{m}$ equal to 40 years for all countries. Different countries differ mainly in the value of the coefficients a and b. Broadly speaking, a determines the "insurance" component of retirement income, because it is independent of the individual's own lifetime earnings, whereas b captures the private returns to one's own lifetime earnings. In this sense, a retirement system with a high ratio of a/b provides high insurance but low incentives for high earnings and vice versa for a low ratio of a/b. Inspecting the coefficients in the table shows that there is a very wide range of variation across countries. Finally, some countries have a ceiling on pensionable income and entitlements, which is also reported in Table A.3.

UI System. The OECD provides data on UI benefits that would be paid to a qualifying person at different points during the unemployment spell: (i) in the first month after the worker becomes unemployed, and (ii) after 5 years of long-term unemployment, which we will refer to as *initial UI* and *final UI* benefits, respectively. An individual with gross earnings y, who has been employed for m years prior to becoming unemployed will receive an initial UI of

$$\Phi(y, m, s) = min\left(1, \frac{m}{\overline{m}^{UI}}\right)\left[a\overline{y} + b\overline{y}^j\right].$$

As before, $\overline{m}^{UI}$ denotes the minimum number of years required to receive full UI benefits, and partial benefits are received if $m < \overline{m}^{UI}$. We set $\overline{m}^{UI}$ to 20 years for all countries. UI benefits are assumed to decline (every year) linearly between the rates provided by the OECD for initial and final UI levels. Some countries also have an upper level of unemployment insurance denoted by $\overline{UI}$ in Table A.4.

Table A.4: Unemployment Insurance Formulas

	a	b	$\overline{UI}$	Ranges of Income
DEN	0.173	0.258		if $y \leq 0.75\bar{y}$
	0.367			if $y > 0.75\bar{y}$
FIN	0.285	0.100		
FRA	0.010	0.392	2.24	
GER	0.091	0.253	0.90	
NET	0.205	0.246		if $y \leq 1.25\bar{y}$
	0.513			if $y > 1.25\bar{y}$
SWE	0.145	0.375		if $y \leq 0.75\bar{y}$
	0.338	0.118		if $0.75\bar{y} < y \leq \bar{y}$
	0.456			if $y > \bar{y}$
UK	0.301			
US	0.045	0.420		if $y \leq \bar{y}$
	0.465			if $y > \bar{y}$

References

Acemoglu, D. (2003): "Cross-Country Inequality Trends," *Economic Journal*, 113(485), F121–F149. 6

Autor, D. H., L. F. Katz, and M. S. Kearney (2008): "Trends in U.S. Wage Inequality: Revising the Revisionists," *Review of Economics and Statistics*, 90(2), 300–323. 38

Baker, M. (1997): "Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings," *Journal of Labor Economics*, 15(2), 338–375. 3

Ben-Porath, Y. (1967): "The Production of Human Capital and the Life Cycle of Earnings," *The Journal of Political Economy*, 75(4), 352–365. 2

Blau, F., and L. Kahn (1996): "International differences in male wage inequality: institutions versus market forces," *Journal of Political Economy*, pp. 791–837. 6

Borghijs, A., S. Ederveen, and R. de Mooij (2003): "European wage coordination: Nightmare or dream to come true? An economic analysis of wage bargaining institutions in the EU," Discussion paper, European Network of Economic Policy Research. 47

Boskin, M. J. (1977): "Notes on the Tax Treatment of Human Capital," Conference on Tax Research 1975, Washington: Dept. Treasury. 9

Bound, J., C. Brown, and N. Mathiowetz (2001): "Measurement error in survey data," in *Handbook of Econometrics*, ed. by J. Heckman, and E. Leamer, chap. 59, pp. 3705–3843. Elsevier. 24

Browning, M., L. P. Hansen, and J. J. Heckman (1999): "Micro data and general equilibrium models," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford. 22

Castañeda, A., J. Díaz-Giménez, and J.-V. Ríos-Rull (2003): "Accounting for the U.S. Earnings and Wealth Inequality," *The Journal of Political Economy*, 111(4), 818–857. 12

Caucutt, E. M., S. Imrohoroglu, and K. B. Kumar (2006): "Does the Progressivity of Income Taxes Matter for Human Capital and Growth?," *Journal of Public Economic Theory*, 8(1), 95–118. 7

Conesa, J. C., and D. Krueger (2006): "On the optimal progressivity of the income tax code," *Journal of Monetary Economics*, 53(7), 1425–1450. 12

Davis, S. J., F. Kubler, and P. Willen (2006): "Borrowing Costs and the Demand for Equity over the Life Cycle," *Review of Economics and Statistics*, 88(2), 348–362. 25

Domeij, D., and M. Floden (2009): "Inequality Trends in Sweden 1978-2004," *Review of Economic Dynamics*, forthcoming. 45

Erosa, A., L. Fuster, and G. Kambourov (2009): "The Heterogeneity and Dynamics of Individual Labor Supply over the Life Cycle: Facts and Theory," Working paper, University of Toronto. 22, 27, 28

Erosa, A., and T. Koreshkova (2007): "Progressive taxation in a dynastic model of human capital," *Journal of Monetary Economics*, 54, 667–685. 7

Gottschalk, P., and M. Joyce (1998): "Cross-national differences in the rise in earnings inequality: market and institutional factors," *Review of Economics and Statistics*, 80(4), 489–502. 6

Gourinchas, P.-O., and J. A. Parker (2002): "Consumption over the Life Cycle," *Econometrica*, 70(1), 47–89. 25

Guoveia, M., and R. P. Strauss (1994): "Effective federal individual income tax functions: An exploratory empirical analysis," *National Tax Journal*, 47(2), 317–39. 12

Guvenen, F. (2007): "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?," *American Economic Review*, 97(3), 687–712. 3, 25

——— (2009): "An Empirical Investigation of Labor Income Processes," *Review of Economic Dynamics*, 12(1), 58–79. 3, 25, 28

Guvenen, F., and B. Kuruscu (2009): "A Quantitative Analysis of the Evolution of the U.S. Wage Distribution: 1970-2000," *NBER Macroeconomics Annual*. 5, 21, 39, 40

Guvenen, F., and A. A. Smith (2009): "Inferring Labor Income Risk from Economic Choices: An Indirect Inference Approach," Working paper, University of Minnesota. 24, 25

Haider, S. J. (2001): "Earnings Instability and Earnings Inequality of Males in the United States: 1967-1991," *Journal of Labor Economics*, 19(4), 799–836. 25

Heathcote, J., K. Storesletten, and G. L. Violante (2008): "The Macroeconomic Implications of Rising Wage Inequality in the United States," NBER Working Papers 14052, National Bureau of Economic Research, Inc. 22, 35

Heckman, J. J. (1976): "A Life-Cycle Model of Earnings, Learning, and Consumption," *Journal of Political Economy*, 84(4), S11–44. 9

Hornstein, A., P. Krusell, and G. Violante (2007): "Technology-Policy Interaction in Frictional Labor-Markets," *Review of Economic Studies*, 74(4), 1089–1124. 6

Huggett, M., G. Ventura, and A. Yaron (2007): "Sources of Lifetime Inequality," NBER Working Paper 13224. 3, 7, 23

Kahn, L. (2000): "Wage Inequality, Collective Bargaining, and Relative Employment from 1985 to 1994: Evidence from Fifteen OECD Countries," *Review of Economics and Statistics*, 82(4), 564–579. 6

King, R. G., and S. Rebelo (1990): "Public Policy and Economic Growth: Developing Neoclassical Implications," *Journal of Political Economy*, 98(5), S126–50. 9

Kitao, S., L. Ljungqvist, and T. Sargent (2008): "A Life Cycle Model of Trans-Atlantic Employment Experiences," Working paper, USC and NYU. 7

Krebs, T. (2003): "Human Capital Risk and Economic Growth*," *Quarterly Journal of Economics*, 118(2), 709–744. 7

Ljungqvist, L., and T. J. Sargent (1998): "The European Unemployment Dilemma," *Journal of Political Economy*, 106(3), 514–550. 6

——— (2008): "Two Questions about European Unemployment," *Econometrica*, 76(1), 1–29. 6

Lucas, Robert E, J. (1990): "Supply-Side Economics: An Analytical Review," *Oxford Economic Papers*, 42(2), 293–316. 9

McDaniel, C. (2007): "Average tax rates on consumption, investment, labor and capital in the OECD 1950-2003," Working paper, Arizona State University. 26

OECD (1986): *The Tax/Benefit Position of Production Workers 1981-1985*. Organisation for Economic Co-Operation and Development, Paris. 51

Ohanian, L., A. Raffo, and R. Rogerson (2006): "Long-Term Changes in Labor Supply and Taxes: Evidence from OECD Countries, 1956-2004," NBER Working Papers 12786, National Bureau of Economic Research, Inc. 6, 9, 11

Prescott, E. C. (2004): "Why do Americans work so much more than Europeans?," *Federal Reserve Bank of Minneapolis Quarterly Review*, (Jul), 2–13. 6, 9, 11, 22

Rebelo, S. (1991): "Long-Run Policy Analysis and Long-Run Growth," *Journal of Political Economy*, 99(3), 500–521. 9

Rogerson, R. (2008): "Structural Transformation and the Deterioration of European Labor Market Outcomes," *Journal of Political Economy*, 116(2), 235–259. 6