Taxation of Human Capital and Cross-country Trends in Wage Inequality

Fatih Guvenen* Burhanettin Kuruscu† Serdar Ozkan‡

April 19, 2009

Very Preliminary and Incomplete

Abstract

Since the 1970's, wage inequality has increased substantially in the U.S. and U.K while changing little in most continential European countries (CEU). This paper studies the role of labor income tax policies in explaining these different trends. We begin by documenting two new empirical facts that link these inequality trends to tax policies. First, we show that countries with a more progressive labor income tax schedule have significantly lower *before-tax* wage inequality at a point in time. Second, progressivity is also negatively correlated with the *rise* of wage inequality during this period. In addition, countries that experienced the smallest rise (or largest fall) in labor hours also had the smallest rise in wage inequality. We next construct a life-cycle model in which individuals decide each period whether to go to school, to work, or to 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 to human capital. Progressive taxation compresses the wage structure, thereby distorting the incentives to accumulate human capital, in turn reducing the cross-sectional dispersion of wages. When this economy experiences skill-biased technical change (SBTC), progressivity also dampens the rise in wage dispersion over time. Furthermore, these effects of progressivity are compounded by differences in average labor income tax rates: the higher taxes in the CEU reduces labor supply (and the benefit of human capital investments), further muting the response to SBTC. Consequently, as in the data, countries with higher taxes and/or progressivity experience a larger fall in hours and a smaller rise in inequality. We find that differences in tax policies can account for 2/3 of the difference in the level of, and 60% of the rise in, wage inequality between US-UK and the CEU since 1980.

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

†UT-Austin; kuruscu@eco.utexas.edu

‡University of Pennsylvania; ozkan@econ.upenn.edu

1 Introduction

What are the determinants of wage inequality in modern economies? How do these determinants interact with technological progress and government policies? Why has wage inequality increased substantially in the United States and the United Kingdon while changing little in many continential European countries since the 1970's? (See Table 1.) 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.

We begin by documenting two new empirical relationships between wage inequality and labor market (tax) policies. First, we show that countries with a more progressive labor income tax schedule have significantly lower *before-tax* wage inequality at different points in time. Second, progressivity is also negatively correlated with the *rise* of wage inequality over time. These findings reveal a close relationship between wage inequality and tax policies that motivates our focus on tax policies in understanding cross-country inequality trends. However, these correlations on their own fall short of providing a quantitative assessment of the importance of the tax structure for wage inequality (e.g., what fraction of cross-country differences in wage inequality can be attributed to differences in tax policies? etc.). For this purpose, we build a model.

Specifically, we construct a life-cycle model that features some important determinants of wages–most notably, human capital accumulation and idiosyncratic shocks. Here is an overview of the framework. Individuals begin life with a fixed endowment of "raw labor" (i.e., strength, health, etc.) and are able to accumulate "human capital" (skills, knowledge, etc.) over the life cycle. Therefore, there are two "factors" that determines an individual's productivity in the labor market. The aggregate production technology takes raw labor and human capital as two separate inputs, and each individual supplies both of these factors of production at competitively determined prices (wages). 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 time. Individuals who invest full-time for a specified number of years become college graduates. We assume that skills are general (i.e., not firm-specific) and labor markets are competitive. As a result, the cost of on-the-job investment will be completely borne by the workers, and firms will adjust the hourly wage rate downward by the fraction of time invested on the job. Thus, the cost of human capital investment is the forgone earnings while individuals are learning new skills.

We introduce two features into this framework. First, we assume that individuals differ in their ability to accumulate human capital. 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. Workers with high ability invest more than others, accepting lower

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

	1978-1982	2001-2005	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	1.33	1.56	0.23
US-UK	**1.21**	**1.42**	**0.21**

wages early on in return for higher wages later in life. Thus, a key source of wage inequality in this model is the *systematic* fanning out of the wage profiles. Recent evidence from panel data on individual wages provide support for individual-specific growth rates in wages; see, for example, Baker (1997), Guvenen [2007, 2009], and Huggett et al. [2007].

The second element in the model, and the main driving force behind the changes in wages during this period, is skill-biased technical change (SBTC), which is modeled as a rise in the productivity of human capital relative to raw labor. Because of the two-factor structure described above, the marginal cost of human capital investment (i.e., forgone earnings) is proportional to the prices of both human capital and raw labor, whereas the marginal benefit of investment is proportional only to the price of human capital. Therefore, SBTC increases the benefit more than the cost, resulting in a rise in human capital investment. Furthermore, the strength of this investment response increases with the ability level, implying that those with high ability increase their investments more than others, accepting even lower wages early in life in return for even higher wages later in life. Therefore, cross-sectional wage inequality rises due to the further fanning out of wage profiles after SBTC.

The model described here provides a central role for policies that compress the wage structure—such as progressive income taxes—because they hamper the incentives to accumulate human capital. This is because progressive income taxes reduce (after-tax) wages at the higher end of the wage distribution while boosting at the lower end. As a result, they reduce the marginal benefit of investment (the higher wages in the future) relative to the cost of investment (the current forgone earnings), and hinder investment.[1] Therefore,

[1]A similar effect is caused by minimum wage laws, which impose an upper bound on the amount of on-the-job human capital investment, by effectively preventing firms from creating jobs that offer low initial wages (below the legal minimum) but higher training opportunities. We also model minimum wage laws in

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

	Total Change	Percentage due to	
	in Log 90-10	Log 90-50	Log 50-10
CEU	0.07	91%	9%
US-UK	0.21	76%	24%
Difference	0.14	68%	32%

individuals in an economy with a compressed wage structure will not increase their investments as much as in an economy with a less progressive tax schedule. As a result, the model predicts that countries with a more redistributive tax system will not experience a large increase in inequality in response to SBTC, but will also not be able to accumulate the requisite human capital, and therefore experience the growth surge that happens several decades after the onset of SBTC.

We assume all countries to 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 (and consumption) tax schedules, and in unemployment insurance and retirement benefits system, among others (although it turns out that the tax system is the dominant factor in the quantitative results below). These policy differences explain about 2/3 of the observed gap in wage inequality between US-UK and CEU at the beginning of 21st century. Second, when we choose the degree of skill-biased technical change between 1980 to 2000 to match the rise in wage inequality in the US during this period, the model explains about 60% of the observed gap in the *rise* in wage inequality between US-UK and CEU during this time period. Finally, the smaller inequality in the CEU comes at the expense of lower employment and lower GDP per capita since workers do not accumulate as much human capital as in the US.

In a recent paper, Guvenen and Kuruscu [2007] has studied a stylized version of the present framework—one that abstracts from idiosyncratic shocks as well as from all the institutional details studied here—and applied it to U.S. data. It concluded that even that stark version of the model provides a fairly successful account of several trends observed in the U.S. data since the 1970's, including the rise in overall wage inequality; the initial fall in the college premium in 1970's and the subsequent strong increase in the next two decades; the rise in within-group wage inequality; the stagnation in aggregate productivity growth, and the small rise in consumption inequality despite the large rise in wage inequality. In this paper, we build on this research by explicitly modeling labor market institutions and allowing for idiosyncratic shocks to provide a detailed quantitative assessment of the role

the paper.

of policies for wage inequality.

The paper is organized as follows. The next section starts with a stylized model to explain the various channels through which the tax structure affects human capital investment and, consequently, inequality. It then discusses how the country-specific tax schedules are estimated and uses the estimates to document some empirical links between taxes and inequality. Section 3 presents the full model and the calibration is carried out in Section 4. Section 5 discusses the quantitative implications. Section 6 concludes.

2 US-UK 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. Although what is meant by progressivity is well-understood as a qualitative concept,[2] the precise empirical measure of progressivity we focus on is suggested by the model studied here. 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 Intuition in A Stylized Model

Consider an individual who derives utility from consumption and leisure and has access to borrowing and saving at a constant interest rate, r. 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 individuals' 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 corresponding marginal tax rate function is denoted by

[2] The American Heritage Dictionary defines progressive taxes as "..increasing in rate as the taxable amount increases." Similarly, the Britannica Concise Encyclopedia defines it as "tax levied at a rate that increases as the quantity subject to taxation increases."

$\tau(y)$. Putting these pieces together, the problem of a type j individual can be written as:

$$
\begin{aligned}
\max_{c_s, a_{s+1}, i_s} \sum_{s=1}^{S} \beta^{s-1} u(c_s, n_s) & \\
s.t & \\
c_s + a_{s+1} &= (1 - \bar{\tau}_n(y_s))y_s + (1 + r)a_s \\
h_{s+1} &= h_s + A^j \left(h_s i_s n_s\right)^{\alpha} \qquad (1)\\
y_s &= P_H h_s (1 - i_s) n_s \qquad (2)
\end{aligned}
$$

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

Inelastic Labor Supply. First, suppose that labor supply is inelastic, in which case $n_s = 1$. The optimality condition for human capital investment is (assuming an interior solution):

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

The left hand side is the marginal cost of investment, whereas 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 when taxes are flat-rate, ie, $\tau'(y) \equiv 0$. In this case, all terms involving taxes cancel out and the FOC reduces to:

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

Thus, flat-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).

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

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

As long as the individual's earnings grow over the lifecycle, all tax ratios on the right hand side will be less than one (because of progressivity), which will depress the marginal benefit of investment, and in turn reduce human capital accumulation. Therefore, these tax ratios provide a key measure of the distortion created by progressive taxes. We refer to each ratio as a *progressivity wedge*. We will also refer to the entire discounted value of future wedges as the *cumulative wedge,* which is a convenient summary measure of total distortion.[3]

To understand the effect of progressive taxes on wage inequality, first note that the distortion created by progressive taxes differs systematically across ability levels. At the low end, individuals with very low ability whose optimal plan involve 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 involves low wage earnings early in life and high earnings later will face very large wedges, which will dampen their investment behavior. Thus, progressivity will reduce the cross-sectional dispersion of human capital, and consequently, the wage inequality in an economy.[4]

Endogenous Labor Supply. Second, consider now the the case with elastic labor supply. The FOC becomes:

$$C_j'(Q_s^j) = \{\beta \frac{1-\tau(y_{s+1})}{1-\tau(y_s)} n_{s+1} + \beta^2 \frac{1-\tau(y_{s+2})}{1-\tau(y_s)} n_{s+2} + ... + \beta^{S-s} \frac{1-\tau(y_S)}{1-\tau(y_s)} n_S\}, \tag{5}$$

where now the marginal benefit accounts for the utilization rate of human capital, which depends on the labor supply choice. Now, once again, let us consider the effect of flat rate taxes. The intra-temporal optimality condition implies that labor supply choice depends on the tax rate and the level of human capital. For example, assuming a separable power utility function we get: $(1 - n_s)^{-\gamma} = P_H h_s (1 - \tau) c_s^{-\sigma}$. Therefore, a higher level of tax 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

[3]Finally, it is easy to see that the redistributiveness of the pension system has an effect that works 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 study both in the full model below.

[4]Notice that the price of human capital P_H does not appear in any version of the FOC displayed above, the implication being that *it has no effect* on the amount of investment undertaken by individuals. This is problematic, especially because our goal is to study the impact of skill-biased technical change that raises the value of human capital. When we move on to the full model below, we are going to consider a different technology for human capital accumulation that circumvents this problem.

Table 3: Tax Function Parameter Estimates

Country	a_0	a_1	a_2	n	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

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, so on and so forth. Therefore, with endogenous labor supply, even a flat-rate tax does have an effect on human capital investment, and this effect can be quite large because of the amplification described here. (Of course, progressive taxes continue to depress human capital investment.) Because average labor hours differ significantly across countries and over time (Prescott (2004), Ohanian, Raffo, and Rogerson (2008)), it is also useful to consider this second measure of wedge that takes into account each country's utilization rate of its human capital in addition to its tax structure. We distinguish this latter wedge in equation (5) with a star: *Progressivity Wedge**.

To summarize, the basic model studied here implies that countries with a more progressive tax system will have a lower (before-tax) wage inequality. Furthermore, it will become clear below that these countries will experience a smaller rise in wage inequality in response to SBTC.

2.2 Deriving the Country-Specific Tax Schedules

For each country we follow the same procedure as described here. The OECD web site provides a tax calculator that estimates the total labor income tax for all income levels between 1/2 of average earnings (hereafter, AW) to two times AW. The calculator 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 (children exemptions, deductions for taxes paid, social assistance, housing assistance, in-work benefits, etc.).[5] Using this tool, we calculate the average income tax

[5]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. Another notably absent component is the social security contributions made by the employer.

rate, $\bar{\tau}(y)$, for 50%, 75%, 100%, 125%, 150%, 175% and 200% of AW. One possible approach would be to approximate these data points with a flexible functional form, which can then serve as the average tax schedule for the relevant country. It turns out however that this approach does not always produce sensible results for the tax schedule for income levels much beyond 200% of AW, which is relevant when individuals solve their dynamic program. Fortunately, there is another piece of information available from OECD that allows us to overcome this difficulty. 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 the appendix, we use this information to generate average tax rates at income levels beyond two times AW. Then we fit the following smooth function to the available data points:

$$\bar{\tau}(y/AW) = a_0 + a_1(y/AW) + a_2(y/AW)^n$$

where AW is the average wage earnings.

The parameters of the estimated average tax functions for all countries are reported in Table 3, 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 in the relevant regions (up to 10 times AW). The lowest R^2 is 0.984 and the mean is 0.991 indicating a fairly good fit. In figure 1 we plot the estimated functions for three countries: the 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 $PW(0.5, 0.5n)$ for $n = 1, 2, .., 6$, which are essentially the wedges faced by an individual who starts life at half the average earnings in that country and looks towards 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 might conjecture. US and 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 for the full-blown model in Section X): for example, a young worker who earns half the average wage today and aims to earn two times the average wage ($n = 4$) in the future loses about 12% of his before-tax wage in the US and UK compared to 26% in Denmark and Finland. These differences grow with the ambition level of the individual, dampening the incentives to accumulate human capital especially at the top of the distribution.

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

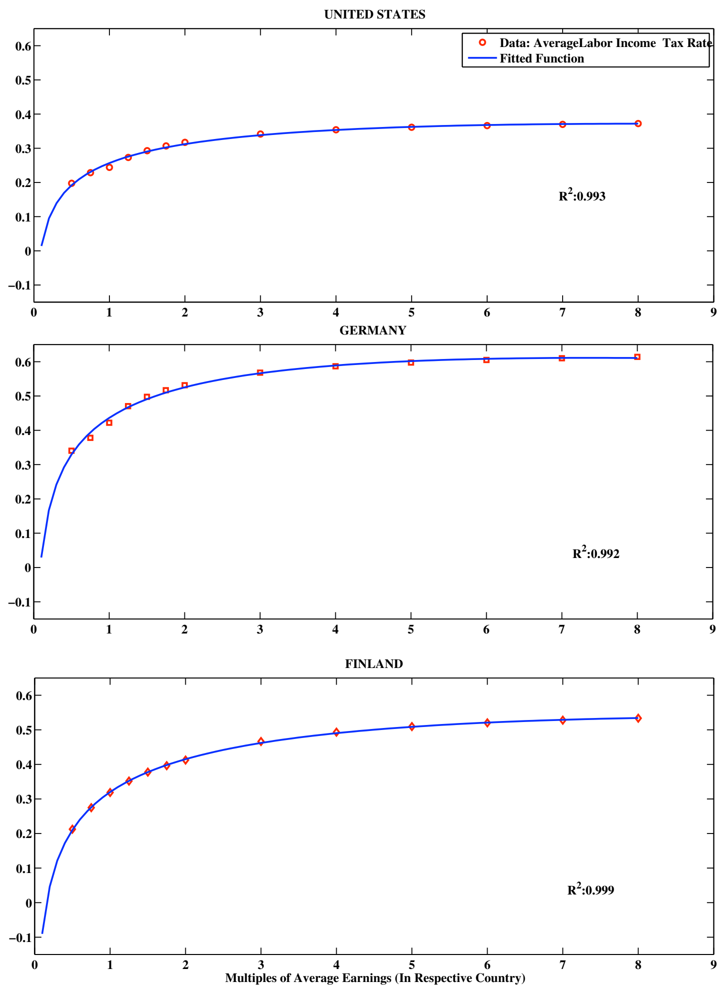
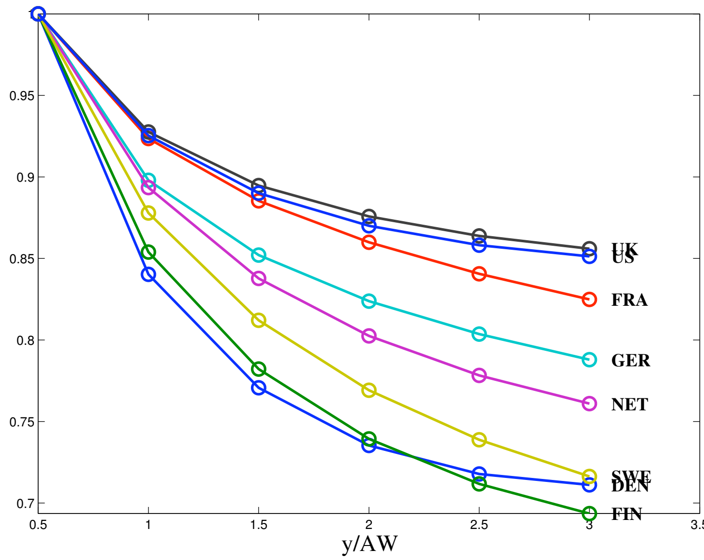


Figure 2: Progressivity Wedges by Income Level: $\frac{1-\tau(n\times 0.5)}{1-\tau(0.5)}$ for $n = 2, 3, .., 6$.



2.3 Taxes and Inequality: Cross-Country Empirical Facts

As explained above, 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 the average wage earnings (hereafter, AW) in each country to 1 and focus on the progressivity wedge between half the average earnings and twice the average earnings: $PW(0.5, 2) = \frac{1-\tau(2)}{1-\tau(0.5)}$. Similarly, the progressivity wedge* is defined as: $PW^* = PW \times \bar{n}$ where $\bar{n}$ is the average hours per person between 2001 and 2005 in the respective country.

The wage inequality data come from the OECD's LFS database and are derived from the gross (i.e., before tax) wage earnings of full-time, full-year (or equivalent) workers.[6] This is the appropriate measure as it more closely corresponds to the marginal product of each worker that is paid to the worker in the model. The fact that the inequality data pertains to before tax wages is important to keep in mind; if it were after-tax wages, the correlation between the progressivity of taxes and inequality would be mechanical and therefore not surprising at all.

Figure 3 plots the relationship between before-tax wage inequality and progressivity wedge in the 2000's. Countries with a higher 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 0.84. Repeating the same

[6]Gross earnings means before taxes and social security contributions. An exception is France, for which earnings are net of employee social security contributions. Other issues regarding data comparability and details are proveided inthe Appendix.

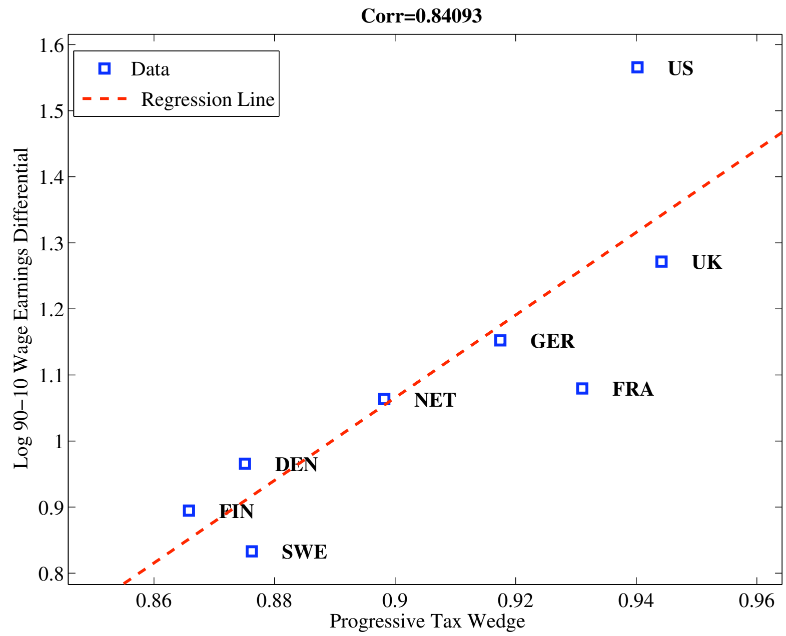
Table 4: Progressivity and Wage Inequality

	$PW(0.5, 2.5)$	$PW^*(0.5, 2.5)$
Wage Inequality:		
2000's:		
90-10	.83	.75
90-50	.84	.73
50-10	.73	.67
1980:		
90-10	.87	.51
90-50	.74	.38
50-10	.88	.57
Change over Time		
90-10	.11	.63
90-50	.39	.86
50-10	-.14	.34

Table 5: Cross-Correlation of $PW(n, m)$ and Log 90-10 Wage Differential

1980's						
n $\downarrow$	$m \rightarrow$	1.0	1.5	2.0	2.5	3.0
0.5		.82	.85	.86	.87	.88
1.0			.87	.87	.87	.86
1.5				.86	.84	.79
2.0					.76	.65
2000's						
0.5		.73	.77	.80	.83	.85
1.0			.81	.84	.87	.89
1.5				.87	.88	.88
2.0					87	.82

Figure 3: Progressivity Wedge and the ***Level*** of Wage Dispersion, 2000's



calculation using the utilization adjusted wedge (PW^*) yields a correlation of 0.75. Both of these relationships are consistent with the simple human capital model with progressive taxes presented above.

We next turn to the change in inequality over time. Figure 4 plots the progressivity wedge* versus the change in the log 90-10 earnings differential. Countries with a more progressive tax system in 2000's have experienced a smaller rise in wage inequality since 1980's. 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 remarkably strong (0.85), whereas the correlation with the 50-10 differential is much weaker (only 0.34; see figures 5 and 6). This result is consistent with the idea that the distortions 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 of the wage distribution where other factors, such as unionization, minimum wage laws, etc. could be more important.

Finally, figure 7 plots the change in hours per person from 1980 to 2003 against the change in wage inequality during the same time period, which shows a fairly strong positive relationship with a correlation of 0.63 for the eight countries in our main sample (and a correlation of 0.57 for all 12 countries on which we have data). One possible explanation for this relationship is suggested by the model described in the previous section (and explored more thoroughly below): with endogenous labor supply, higher (and/or more progressive)

Figure 4: Progressivity Wedge* and ***Change*** in Log **90-10** Diff.: 1978 to 2005

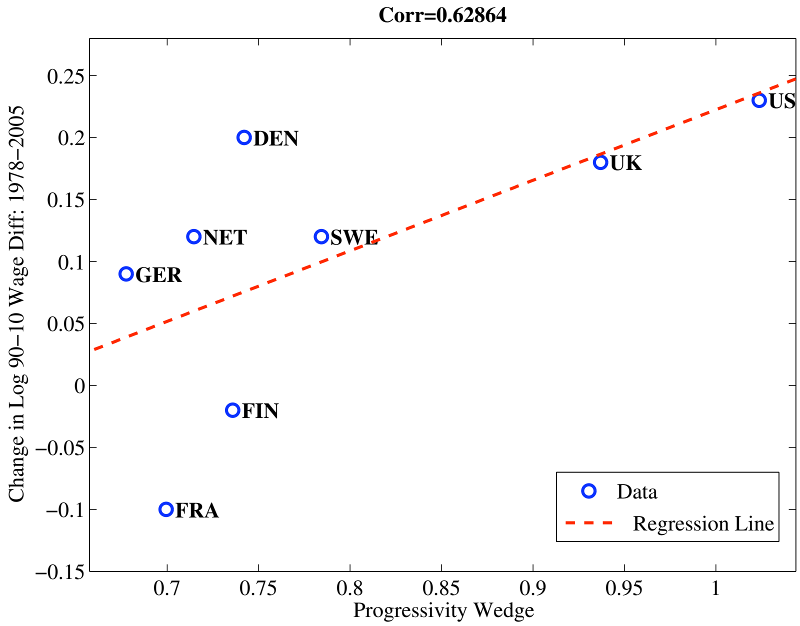


Figure 5: Progressivity Wedge* and ***Change*** in Log **90-50** Diff.: 1978 to 2005

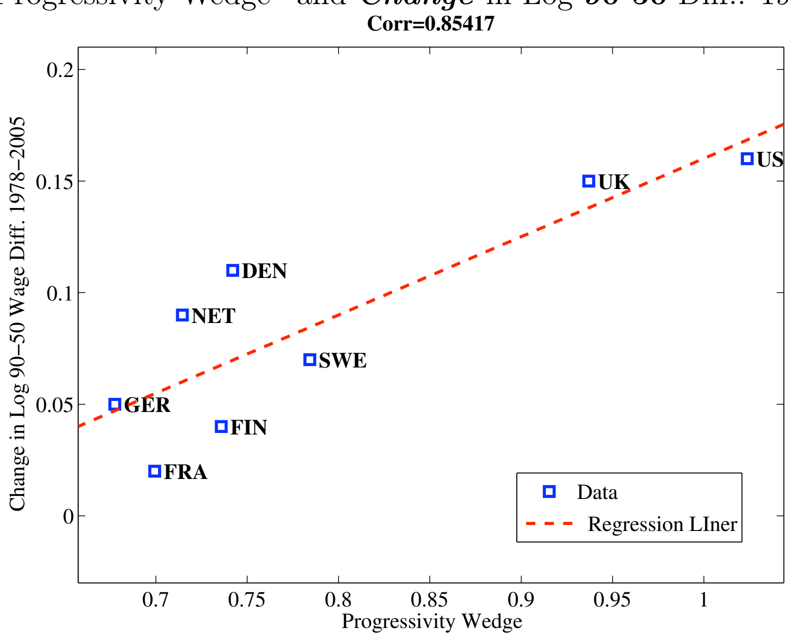
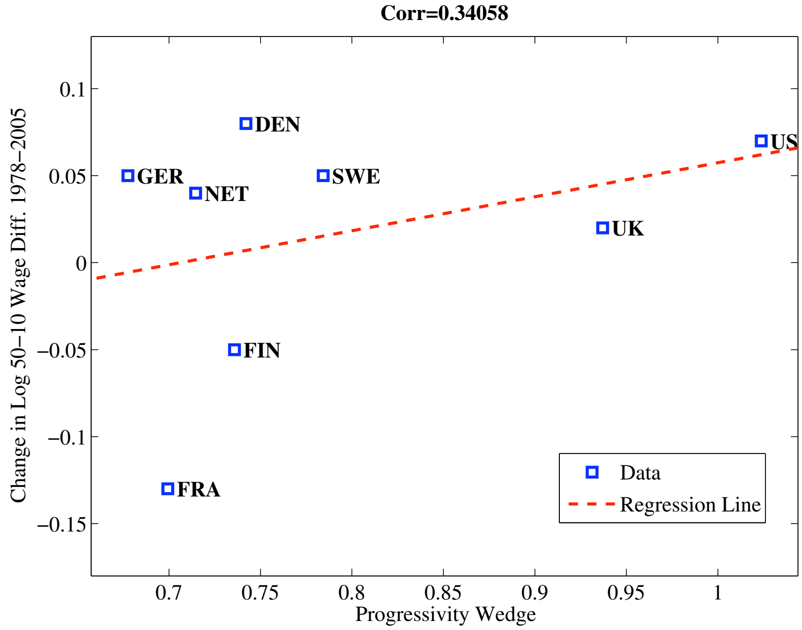


Figure 6: Progressivity Wedge* and ***Change*** in Log **50-10** Diff.: 1978 to 2005



labor income taxes depress labor supply, which in turn dampens human capital accumulation and wage inequality. Consequently, countries that experience a smaller increase (or larger decline) in their labor supply are also those who experience a smaller rise in wage inequality, as seen in this figure.

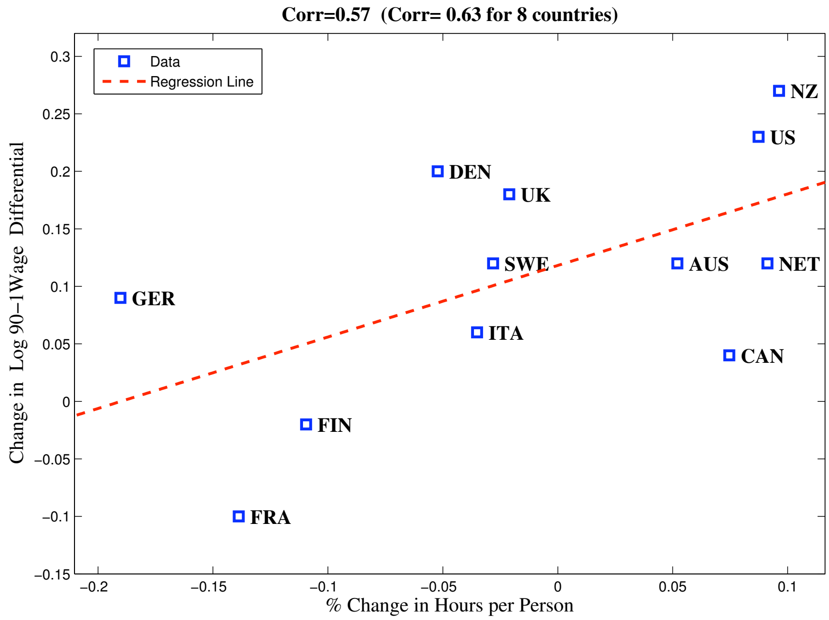
3 The Model

The model we use for the quantitative 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 the common separable power utility form:

$$u(c,n)=\frac{c^{1-\sigma}}{1-\sigma}+\psi\frac{(1-n)^{1-\gamma}}{1-\gamma}. \tag{6}$$

If an individual chooses to work, as before, she can allocate a fraction of her working hours, i, to human capital investment. However, more realistically, we now assume that $i\in[0,\chi]$ where $\chi<1$. An upper bound less than 100 percent 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 due to minimum wage laws. Individuals *can* invest

Figure 7: Change in Hours/Person and in Log 90-10 Wage Diff. Between 1980 and 2000



full-time by attending school ($i = 1$) and enjoy leisure for the rest of the time. Thus, the choice set for investment time is: $i \in [0, \chi] \cup \{1\}$, which is non-convex when $\chi < 1$.

An individual may choose to be unemployed at age s, $n_s = 0$, in which case she receives unemployment benefit payments as specified below. Finally, individuals retire at age R and receive a pension that depends on their pre-retirement earnings as well as their year of service, again, in a manner that mimics the pension system in the relevant country. Everybody dies at age T. Each component of the model is now described in further detail.

3.1 Human Capital Accumulation

Individuals begin life with an endowment of "raw labor" (i.e., strength, health, etc.), which is constant over the life cycle, and are able to accumulate "human capital" (skills, knowledge, etc.) over the life cycle. There is a continuum of individuals in every cohort, indexed by $j \in [0, 1]$, who differ in their ability to accumulate human capital, denoted by A^j (also referred to as their "type"). Below we suppress the superscript j when it does not create confusion. Let l denote raw labor and h^j_s denote the human capital of an s-year-old individual of type j. 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 (potentially stochastic) wage rates, denoted by P_L and P_H, respectively.[7]

Individuals begin their life with zero human capital and each period produce new human capital, Q^j, according to the following technology:

$$Q^j = A^j \left((\theta_L l + \theta_H h^j) i^j n^j \right)^{\alpha} \tag{7}$$

where i^j is the fraction of time devoted to human capital investment, henceforth referred to as "investment time"; and Q^j is the newly produced human capital which will be referred to simply as "investment" in the rest of the paper. According to this formulation new human capital is produced by combining the existing stocks of raw labor and human capital with the available investment time. A key parameter in this specification is A^j, which determines the productivity of learning. Due to the heterogeneity in A^j, individuals will differ systematically in the amount of investment they undertake, and consequently, in the growth rate of their wages over the life cycle. Another important parameter is $\alpha \in [0, 1]$, which determines the degree of diminishing marginal returns in the human capital production function. Finally, both raw labor and human capital depreciates every period and we assume that this happens at the same rate δ.

3.2 Idiosyncratic Shocks and Earnings

Individuals receive idiosyncratic shocks to the efficiency of the labor they supply to the firm. 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 $\Pi(\epsilon' \mid \epsilon)$ which is identical across agents and over the lifecycle. The observed *total wage income* of an individual who receives a shock ϵ and spends i_s^j fraction of his time learning new skills is given by

$$y_s^j \equiv \epsilon \left[P_L l + P_H h_s^j \right] n_s^j (1 - i_s^j) = \underbrace{\epsilon \left(P_L l + P_H h_s^j \right) n_s^j}_{\text{Potential earnings}} - \underbrace{\epsilon \left(P_L l + P_H h_s^j \right) n_s^j i_s^j}_{\text{Cost of investment}}, \tag{8}$$

where $\left[P_L \epsilon l + P_H \epsilon h_s^j \right] n_s^j$ is the "potential earnings" of an individual—that is, the income an individual would earn if he spent all his time on the job producing for his employer. Therefore, wage income can be written as the potential earnings minus the "cost of investment,"

[7] The structure we have in mind is not one where an individual works at manual tasks (using raw labor only) some fraction of the time and at cognitive (or skill-intensive) tasks at other times. Instead the worker employs both factors of production simultaneously in producing output. For example, a college professor uses both his/her body and knowledge/skills at the same time when teaching, although probably at different proportions than a farmer, an auto mechanic, or a brain surgeon.

which is simply the forgone earnings while individuals are learning new skills. Finally, the *hourly wage* rate is given by

$$\begin{aligned} w_s^j = y_s^j / n_s^j &= \epsilon \left(P_L l + P_H h_s^j \right) - \epsilon \left(P_L l + P_H h_s^j \right) i_s^j \\ &= \epsilon \left(P_L l + P_H h_s^j \right) - \epsilon C(Q_s^j)/n_s^j \end{aligned}$$

where $C(Q_s^j) \equiv (P_L l + P_H h_s^j)\, n_s^j i_s^j$ is the opportunity cost of investment for this model.

3.3 Aggregate Production Function

Let H and L denote the aggregate amounts of human capital and raw labor used in production–ie, net of the time allocated to learning new skills–at a point in time. Because all shocks are idiosyncratic, they have no effect at the aggregate level. An aggregate firm uses these two inputs to produce a single good, denoted by Y, according to the following CES production function:

$$Y = Z \left([\theta_L L]^\rho + [\theta_H H]^\rho \right)^{1/\rho}, \tag{9}$$

where $\rho \leq 1$, and Z is the total factor productivity (TFP). We assume that physical capital is not used in production. Notice that human capital and raw labor enter the aggregate production function and human capital function with the same weights (compare equations (7) and (9)). In Guvenen and Kuruscu [2007], we argue that this specification produces several plausible implications for the behavior of wages in addition to simplifying the solution of the model.

The firm maximizes profits ($\equiv Y - P_L L - P_H H$) period by period. The first order conditions from this maximization problem can be rearranged to obtain the price of human capital *relative* to raw labor:

$$\frac{P_H}{P_L} = \left(\frac{\theta_H}{\theta_L} \right)^\rho \left(\frac{H}{L} \right)^{\rho - 1}. \tag{10}$$

While the aggregate production function has the same CES form commonly used in the literature, its inputs are different than what is typically assumed (eg., Katz and Murphy [1992]). In most previous work, H and L denote the labor supplied by workers with college and high school education respectively. Therefore, a change in the price of H relative to L has the same effect on all individuals within an education group. As a result, there is no within-group inequality and the total wage dispersion is the same as the college-high school premium. Of course, empirically, the latter explains less than 1/3 of total inequality, which is a problem if the goal is to study the evolution of total wage inequality as we do here. In contrast, in the present model, all workers have some endowment of human capital (which varies by ability and age) and l (which is the same for all), and every worker contributes to both factors of production. Therefore, a change in the price of H relative to L affects all individuals differently depending on their ability level as well as their age. Below, we consider the special case of $\rho = 1$, which implies $P_H/P_L = \theta_H/\theta_L$.

3.4 Government: Taxes and Transfers

3.4.1 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. Specifically, if a worker becomes unemployed at age s, the initial level of the unemployment payment she receives increases with her years of work before becoming unemployed, denoted by m, and also varies (typically decreases) with the duration of the unemployment spell. The precise nature of the dependence on past service and unemployment spell varies by country, which will be incorporated during calibration. Finally, in most countries the replacement *rate* falls with the level of pre-unemployment income, which is also captured here. $\Phi(y^*, m, s)$ denotes the unemployment benefit function of an s year old individual who has worked a total of m years before becoming unemployed (at age $s - m$). Although, in reality, unemployment payments depend on the pre-unemployment earnings, making this dependence explicit will introduce (some measure of) past earnings as an additional state variable, which complicates the solution of the 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 if he did not have the option of receiving unemployment insurance. For the precise problem that yields y^*, see the appendix.

After retirement individuals receive constant pension payments every period. The amount of these payments increase with the years of service completed up to age R, as well as with wage earnings at retirement age. The precise nature of these payment schedules are modeled to mimic country-specific features. The pension function is denoted as $\Omega(y_R, m)$.

3.4.2 The Tax System and the Government Budget

The government imposes a flat-rate ($\bar{\tau}_c$) consumption tax as well as a potentially progressive labor income tax, $\bar{\tau}_n(y)$.[8] These receipts are used for three purposes: (i) finance the benefits

[8]We ignore capital income taxes for now.

system, (ii) finance government expenditure, G, that does not yield any direct utility to consumers (either due to corruption or waste), and (iii) the residual budget surplus or deficit is distributed in a lump-sum fashion, denoted *Tr*, to all households regardless of employment status.

3.5 Market Structure

Individuals trade a full set of one-period Arrow securities $a(\epsilon' \mid \epsilon)$ that pays one unit of consumption good next period conditional on shock being ϵ' next period and today's shock being ϵ. The price of each Arrow security is denoted by $q(\epsilon' \mid \epsilon)$.

3.6 Individuals' Dynamic Program

Individuals cannot accumulate human capital while unemployed. (Notice that an individual who is enrolled in school is *not* considered unemployed in this model.) The problem of an s-year old individual with ability A^j (we suppress ability type for clarity), who has worked for m years, entered the period with h units of human capital and a units of Arrow securities and has observed shock ϵ is given by:

$$
\begin{aligned}
V(h,a,m;\epsilon,s) &= \max_{c,n,i,a'(\epsilon')} \left[u(c,n) + \beta \sum_{\epsilon'} \Pi(\epsilon' \mid \epsilon) V(h', a'(\epsilon'), m'; \epsilon', s+1) \right] \\
s.t & \\
(1+\bar{\tau}_c)c + \sum_{\epsilon'} q(\epsilon' \mid \epsilon) a'(\epsilon') &= (1-\bar{\tau}_n(y))y + a + Tr \\
y &= \underbrace{[\epsilon(P_L l + P_H h)(1-i)]}_{w} n I_n + \underbrace{\Phi(y^*, m, s)}_{\text{UI Benefits}}(1-I_n) \qquad (11) \\
h' &= (1-\delta)h + A([\theta_H h + \theta_L l]ni)^{\alpha}, \qquad l' = (1-\delta)l \\
m' &= m + I_n \\
i &\in [0,\chi] \cup \{1\}
\end{aligned}
$$

where I_n is an indicator function that is equal to 1 if the agent is working in that period and 0 is he is unemployed; and y^* is defined above.

The solution to problem (11) gives a set of functions $V(h,a,m;\epsilon,s)$, $c(h,a,m;\epsilon,s)$, $a'(h,a,m;\epsilon',\epsilon,s)$, $n(h,a,m;\epsilon,s)$, $m'(h,a,m;\epsilon,s)$, $Q(h,a,m;\epsilon,s)$, and implied wage earnings $y(h,a,m;\epsilon,s)$. We will need these expressions in the definition of equilibrium.

After retirement, individuals receive a pension and there is no human capital investment. The retirement pension depends on the earnings one period before retirement. Since markets are complete and 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

$$W^R(a, y_{R-1}, m; s) = \max_{c,a'} \left[u(c, \bar{n}) + \beta W^R(a', y_{R-1}, m; s+1)\right] \tag{12}$$
$$s.t$$
$$(1 + \bar{\tau}_c)c + qa' = (1 - \bar{\tau}_n(y_s))y_s + a + Tr$$
$$y_s = \Omega(y_{R-1}, m)$$

Since the pre-retirement earnings y_{R-1} depends on the state of the individual at age $R-1$, i.e. it is given by $y_{R-1} = y(h, a, m; \epsilon, , R-1)$, the problem above implicitly depends on pre-retirement state $(\epsilon, a, h, m; R-1)$. However, we do not need to keep track of all the variables since y_{R-1} and m are the only variables that matters for pension.[9] However, in the definition of equilibrium, we need to identify the fraction of agents in state $(h, a, m; \epsilon, R-1)$ to determine the total social security payments.

Definition 1 *The stationary equilibrium for this economy is a set of equilibrium decisions rules $c(x)$, $n(x)$, $Q(x)$, $i(x)$ and $a'(\epsilon', x)$; and value function functions for $V(x)$ 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); and a time-invariant measure $\Lambda(x)$ such that*

1. *Given prices P_H and P_L, the labor income tax function $\bar{\tau}(y)$, consumption tax $\bar{\tau}_c$ and government policy functions Φ and Ω; individuals solve problems in (11) and (12).*

2. *Prices are determined competitively.*

3. *Aggregate amount of raw labor and human capital are given by*

$$H = \int_x \epsilon h(x)(1 - i(x))d\Lambda(x)$$
$$L = \int_x \epsilon l(1 - i(x))d\Lambda(x)$$

[9] Basically, the value function of the agent the year before retirement is $V(\epsilon, a, h, m; R-1) = \max_{c,n,Q,a'(\epsilon')} u(c, n) + \beta E W^R(a'(\epsilon'), y(\epsilon, a, h, m; R-1), m; R)$ subject to constraints of problem (11). It should be clear that individuals choose the same amount of Arrow securities for all states (which is equivalent to holding a riskless bond) since there is no uncertainty in future income.

4. The government budget balances:

$$\int_{x,s\leq R} \bar{\tau}_n(y(x))y(x)I(n(x)>0)d\Lambda(x) + \int_x \bar{\tau}_c c(x)d\Lambda(x) = G + Tr$$
$$+ \int_{x,s\leq R} \Phi(y^*(x),m,s)I(n(x)=0)d\Lambda(x)$$
$$+ \int_{s>R}\int_x \Omega(y(h,a,m;\epsilon,R-1,j),m)d\Lambda(h,a,m;\epsilon,R-1,j)$$

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.[10] Similarly, the second term is the total tax revenue from the consumption tax, but it is collected from all agents includding the retirees. On the right hand side, the pension payments only depend on individuals' last period income before retirement. Thus, all individuals who are in the same state in the last period before retirement will earn the same retirement income. As a result, total amount of pension payments in the economy is calculated by integrating all agents at age $R-1$.

3.7 Optimal Investment Decision

Before moving on to the quantitative analysis, it is useful to examine the effect of skill prices on investment behavior in this full model, which features a two-factor structure (with raw labor and human capital) differently than the simple model in section 2.1. Recall that in that model a rise in the price of human capital had no effect on investment behavior. Turning to the full model, a similar first order optimality condition can be derived by 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. While these features are important for the quantitative results, useful insights can be gained without them. Under these assumptions, the first order condition is:

$$C_j'(Q_s^j) = \theta_H\{\beta\frac{1-\tau(w_{s+1})}{1-\tau(w_s)}n_{s+1} + \beta^2\frac{1-\tau(w_{s+2})}{1-\tau(w_s)}n_{s+2} + ... + \beta^{S-s}\frac{1-\tau(w_S)}{1-\tau(w_s)}n_S\} \tag{13}$$

The key observation is that optimal investment, Q_s^j, only depends on the level of θ_H—not on the levels of θ_L or Z. This is because the opportunity cost of investment depends on

[10]The pension and unemployment benefit functions are written as paying after-tax income so they are not taxed again.

the prices of both raw labor and human capital (see equation (2)), whereas the marginal benefit is only proportional to the price of human capital. As a result, a higher level of θ_H (for example, due to SBTC) increases the marginal benefit more than the marginal cost, resulting in higher investment. This feature is an important difference between the current framework and the standard Ben-Porath model (studied in section 2.1. Compare (13) to (4) and (5). In the latter, a higher price of human capital (which is the only factor of production since there is no raw labor) affects the cost and benefit of investment exactly the same way, leaving the trade-off—and therefore the investment decision—unaffected. It is precisely for this reason that it is difficult to think of the concept of "returns-to-skill" in that framework, because a higher price of human capital has no effect on the decision to invest. Instead in the present model θ_H/θ_L is a measure of returns-to-skill, and affects investment in human capital without necessarily implying anything about aggregate productivity (which is captured by Z and also has no effect on investment incentives for the same reason discussed for the Ben-Porath model).

Using (13), optimal investment choice can be solved for explicitly:

$$Q_s^j = \left(A^j\right)^{1/(1-\alpha)} \left[\alpha MB\right]^{\alpha/(1-\alpha)}. \tag{14}$$

This expression highlights the main sources of heterogeneity in this model: (i) individuals with higher learning ability invest more in human capital: $\partial Q_s^j/\partial A^j > 0$; (ii) more importantly, their investment responds more strongly to SBTC: $\partial^2 Q_s^j/\partial\theta_H \partial A^j > 0$; (iii) investment goes down over the life cycle: $\partial Q_s^j/\partial s < 0$; and finally, younger individuals respond more strongly to SBTC: $\partial^2 Q_s^j/\partial\theta_H \partial s < 0$.

4 Quantitative Analysis

4.1 Calibration [subsection to be revised]

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. We then assume that other countries share the same parameter values with the US along unobservable dimensions (such as the distributions of true ability and raw labow), but differ in several dimensions of their labor market policies that are feasible to model and calibrate (specifically, consumption and labor income tax schedules, retirement pension system, unemployment insurance system, and finally legal minimum wages). We will then examine the differences in economic outcomes—specifically in wage dispersion, output, and labor supply—that are generated by these policy differences alone. When we compare economic outcomes over time, we again calibrate the change to match the US data and compare outcomes implied for other countries to the data.

As noted earlier, the present model shares some common features with the framework studied in Guvenen and Kuruscu [2007], where we discuss in detail the justifications for the parameter choices for the US. Below we refer the reader to that paper for more details for the choices of these common parameter values. Other aspects of the calibration specific to the present model are discussed here in more detail.

Individuals enter the economy at age 20 and retire at 65 ($S = 45$). Everybody dies at age 85. The net interest rate, r, is set equal to 5%, and the subjective time discount rate is set to $\beta = 1/(1+r)$. The growth rate of neutral technology level, Z, is set equal to 1.5 percent per year. However, *measured* TFP growth will be different than this number when the amount of investment on-the-job changes over time. We take the curvature of the aggregate production technology, ρ, to be unity implying a linear production function. The motivation for this choice can be found in Guvenen and Kuruscu [2007] where we also study the case with imperfect substitution. The curvature of the human capital accumulation function, α, is set equal to 0.80 broadly consistent with the existing empirical evidence. Higher values of this parameter has sometimes been estimated in the literature (cf. Heckman (1976), Heckman et al. (1998), Kuruscu (2006)). Such values generate a stronger impact of human capital investments on variables of interest (and typically improve the performance of the model).

The remaining parameters of the model are chosen to match some key empirical targets in the US data during the period 2001-2005. First, the weights in the production function, θ_L and θ_H, always appear multiplicatively with raw labor and human capital, so the initial values of these parameters serve only as a normalization (given that H and L are also calibrated separately below). Therefore, we normalize $\theta_L + \theta_H = 1$ and set $\theta_L = \theta_H = 0.5$.

Accounting for Idiosyncratic Shocks. For a meaningful comparison of the model to the data, it is important to account for the fact that the model abstracts from idiosyncratic shocks, which are clearly present in the data. To this end, we assume that the logarithm of the observed wage in the data can be written as

$$\log \widetilde{w}_{s,t}^j = \log w_{s,t}^j + \upsilon_{s,t}^j + \xi_{s,t}^j, \tag{15}$$

where $w_{s,t}^j$ denotes the systematic (or life-cycle) component of wages, and is given by the baseline human capital model in this paper; $\upsilon_{s,t}^j$ represents a first-order autoregressive shock process, and $\xi_{s,t}^j$ is a transitory disturbance with variance σ_ξ^2. Both the innovation to the

AR(1) process and $\xi^j_{s,t}$ are *i.i.d.* conditional on all individual characteristics (including s and A^j). This specification is similar to the econometric processes for wages commonly used in the literature.[11] The key assumption we make is that the variances of these idiosyncratic shocks have been *stationary* during the period under study. Under this assumption, and letting $var(\cdot)$ denote the cross-sectional variance of a variable, we have:

$$var\left(\log \widetilde{w}^j_{s,t}\right) = var\left(\log w^j_{s,t}\right) + \sigma^2_v + \sigma^2_\xi.$$

where σ^2_v denotes the cross-sectional variance of the AR(1) process across all age and ability groups. Two points are easily noted from this expression. First, the *level* of the variance of wages in the model needs to be adjusted by $\left(\sigma^2_v + \sigma^2_\xi\right)$ before it can be compared to the data. Second, the *change* over time in the variance of observed wages will mirror that in the systematic component $\left(\Delta var\left(\widetilde{w}^i_{s,t}\right) = \Delta var\left(w^i_{s,t}\right)\right)$ which allows a direct comparison of the *trend* in the model variances to its empirical counterpart.

Similarly, the implications of the specification in (15) for the first moment of wages can also be seen easily: the average of observed log wages equals that of the systematic component, $E\left(\log \widetilde{w}^j_{s,t}|I\right) = E\left(\log w^j_{s,t}|I\right)$, where I denotes a set of individuals—for example, those in the same age or education group. Therefore, both the level of, and the change in, the first moments of log wages in the model can be directly compared to the data.

4.1.1 Calibrating Model-Specific Parameters

Legal minimum wage.

Distributions of Ability and Raw Labor. Learning ability, A^j, is assumed to be uniformly distributed in the population with the same parameters for every country. As for the calibration of individuals' raw labor endowment, note that the present model is interpreted as applying to human capital accumulation after secondary school. But then, assuming that individuals start out with the same human capital level—may be too restrictive because it seems likely that different individuals would have accumulated different amounts of human capital by the time they make the college enrollment decision. A simple way to model this heterogeneity is by assuming that the amount of raw labor, l, has a non-degenerate distribution in the population. We also assume l to have a uniform distribution that is the same

[11] One caveat of this specification should be noted. Because the idiosyncratic shocks introduced here are multiplicative with $w^j_{s,t}$, it can be shown that if individuals take the existence of these shocks into account when making their investment decision, this would lead to a different optimal choice than the one that generated $w^j_{s,t}$. Although it is possible to modify the human capital problem and solve it in the presence of these idiosyncratic shocks, such an extension comes at considerable computational cost, so we do not tackle this potential complication here.

for all cohorts. Each distribution is fully characterized by two parameters, giving us four parameters to be calibrated.[12] The mean value of raw labor, $E\left[l^j\right]$, is a scaling parameter and is normalized to one, leaving three parameters: (i) the cross-sectional standard deviation of raw labor, $\sigma\left(l^j\right)$, (ii) the mean learning ability, $E\left[A^j\right]$, and (iii) the dispersion in the ability to learn, $\sigma\left(A^j\right)$. These are chosen to match the following three moments:

1. the average cross-sectional variance of log wages between 1965 and 1969,
2. the average level of the log college premium between 1965 and 1969,
3. the mean log wage growth over the *life cycle.*

As discussed above, we need an estimate of the variances σ_υ^2 and σ_ξ^2 to obtain the target value for the cross-sectional wage inequality. Note that, for consistency, these estimates must be obtained from empirical studies that allow for heterogeneity in wage growth rates as implied by the human capital model in this paper.[13] Guvenen (2005) estimates such a specification and reports σ_ξ^2 to be 0.047. Similarly, σ_υ^2 can be calculated to be 0.088 using the estimates in that paper (Table 1, row 2). The average cross-sectional variance of log wages in the U.S. data between 1965 and 1969 is 0.239, implying a target value for the first moment in the model $\left(var\left(w_{s,t}^j\right)\right)$ of 0.104. Second, the log college premium in the U.S. data averaged 0.381 between 1965 and 1969 (and does not require any adjustments), which is the second empirical target we choose. Third, and finally, our target for mean log wage growth between ages 20 and 55 is 50 percent for a cohort of individuals who retire before 1970. 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 (which typically report estimates between 40 and 65 percent; see, for example, Gourinchas and Parker (2002), Davis, Kubler and Willen (2002), Guvenen (2005)).[14]

[12]Notice that we also need to calibrate the cross-sectional correlation of l and A. Since we interpret the heterogeneity in l as arising from investments made prior to college and high-ability individuals are likely to have invested more even before college, it seems reasonable to conjecture that A and l will be positively correlated. Indeed, Huggett et al (2006b) estimate the parameters of the standard Ben-Porath model from individual wage data allowing for heterogeneity in A and l, and provide evidence that the two are strongly positively correlated (corr: 0.792). For simplicity we assume perfect correlation between the two. Furthermore, it will become clear later that the heterogeneity in l does not play a significant role in this model, implying that the choice of perfect correlation is not likely to be critical.

[13]This requirement eliminates several well-known empirical papers, such as MaCurdy (1982), Abowd and Card (1989), and Meghir and Pistaferri (2004), among others, which restrict wage growth rates to be the same across the population.

[14]Ideally, we would like to use an estimate of average life-cycle wage growth during the period before 1970 (before SBTC), whereas PSID is only available starting 1968 on. However, we are not aware of any study that estimates the life-cycle (*not* cross-sectional) wage profiles using data from earlier periods. Our calibration implies a mean log wage growth of 62 percent for the cohort that enters the economy in 1968, consistent with the numbers found by these studies during the same period.

Table 6: Baseline Parameterization

Parameter		Value
r	Interest rate	0.05
β	Time discount rate	$1/(1+r)$
α	Curvature of human capital function	0.80
S	Years spent in the labor market	45
ρ	Curvature of aggregate prod function	1.0
χ	Maximum investment time on the job	0.50
$\Delta \log Z$	Growth rate of neutral technology	.015
$E\left[l^j\right]$	Average labor endowment (scaling)	1.0
Parameters calibrated to match 1980 targets:		
$E\left[A^j\right]$	Average ability	.071
$\sigma\left(l^j\right)/E\left[l^j\right]$	Coeff. of variation of labor endowment	.0503
$\sigma\left[A^j\right]/E\left[A^j\right]$	Coeff. of variation of ability	.245
Parameter calibrated to match 2000 var(log(W)) in US		
$\Delta \log\left(\theta_H/\theta_L\right)$	Change in skill-bias from 1980	23%

Table 6 displays the implied values for the distributions of A^j and l^j. Notice that the coefficient of variation of ability is more than four times that of raw labor. Overall, heterogeneity in l has a much more modest effect on the quantitative results than does the heterogeneity in ability. Finally, it should be stressed that with this calibration the model also matches the cross-sectional variance of wage growth rates observed in the U.S. data (see the discussion above, in section 2.5).

4.1.2 Calibration of the Benefits System, etc

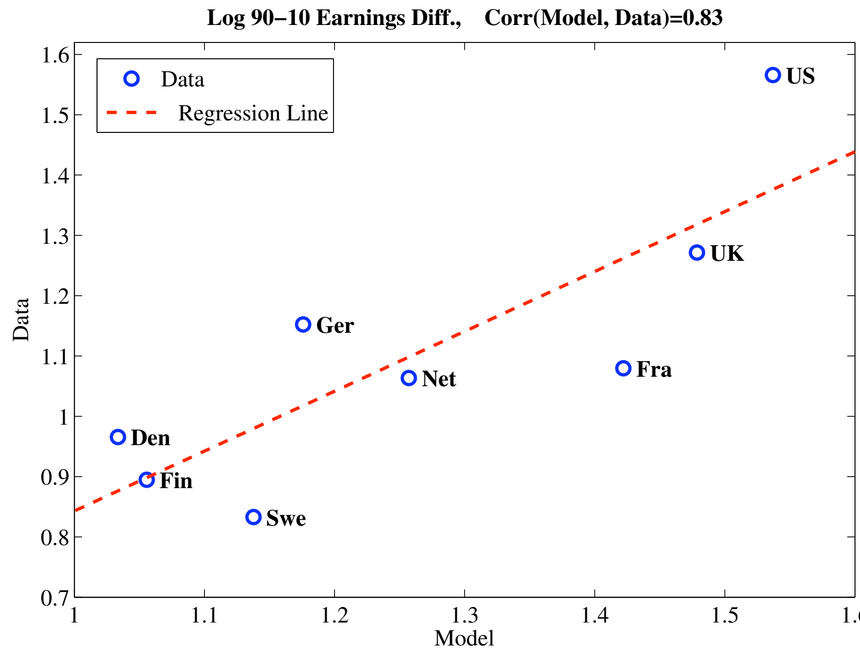
Consumption taxes: Tax rates on consumption are taken from McDaniel (2007). This paper provides average consumption tax rates between 1950 and 2003 for 15 OECD countries. The methodology used here to calculate the tax rate is to divide tax revenue from consumption expenditures by the amount of corresponding expenditure.[15]

5 Results

In this section, we discuss the implications of the calibrated model for the cross-country wage inequality at a point in time, as well as for the rise of inequality over time. First, figure 8 plots the log 90-10 wage differential for each country in the data against the predicted

[15] This is the same methodology used by Mendoza et al (1994).

Figure 8: Wage Dispersion: Model versus Data



value for the same variable by the model. The correlation between the simulated and actual data is 0.83, suggesting that the model is able to capture the relative ranking of countries' inequality in the data.

However, this figure on its own does not allow us to *quantify* how important taxation is for cross-country differences in inequality. For this we turn to table 7. The first column replicates the information displayed before in the third column of Table 1. The second column expresses wage inequality (log 90-10 differential) in each country as a fraction of inequality in the US. The third and fourth columns display the corresponding statistics implied by the calibrated model. For example, in Denmark the actual log 90-10 differential is 0.97, which is approximately 62% of the same variable in the US in 2003. The model generates a log 90-10 differential of 0.89 for Denmark which is 68% of the corresponding figure implied by the model for the US. Similar comparisons show that the model does quite well in explaining the level of wage inequality in Germany (73% of US inequality in the data versus 72% in the model) and does very poorly in explaining the UK (82% in the data versus 97% in the model). The next colum (e) restates the described comparison in an easier to read fashion: i.e., what fraction of the difference between the US and each country is explained by the model. The fraction explained ranges from 36% for France to 103% for Germany. Averaging this figure across all CEU countries show that the model explains about 65% of the actual gap in inequality between the US and CEU in 2003.

Table 7: Quantifying the Contribution of Taxes to Wage Inequality, 2003

	Data		Model		% explained	%explained
	Level	% of US	Level	% of US	(1-d)/(1-b)	inelastic labor
	(a)	(b)	(c)	(d)	(e)	(f)
Denmark	0.97	0.62	0.89	0.68	84%	37%
Finland	0.89	0.57	0.95	0.73	63%	35%
France	1.08	0.69	1.16	0.89	36%	17%
Germany	1.15	0.73	0.94	0.72	103%	29%
Netherlands	1.06	0.68	1.10	0.85	47%	15%
Sweden	0.83	0.54	0.97	0.75	54%	26%
CEU	1.00	**0.70**	1.00	**0.77**	**65%**	**25%**
UK	1.27	0.82	1.27	0.97	13%	3%
US	1.56	1.00	1.30	1.00	–	
US-UK	1.42		1.29			

Given the low Frisch elasticity we used in calibrating the model, it is natural to ask how important endogenous labor supply is for these results. The last column reports the results of calibrating the baseline model with inelastic labor supply (ie, zero Frisch elasticity). The model is calibrated to the same targets as before. Comparing the fraction of wage inequality gap explained in this case to column (e) reveals how important endogenous labor supply is. The model explains a much more modest fraction—25% for all of CEU compared to 65% previously—of wage inequality differences between US and CEU. Therefore, the amplification channel described in Section 2.1 is indeed a powerful source for creating differences in wage inequality levels across countries.

Finally, we turn to the rise in wage inequality over time in response to SBTC. To this end, we choose the skill-bias of technology, θ_H/θ_L, in 2003 such that the model matches (approximately) the total rise of 0.21 in the log 90–10 differential in US-UK from 1980 to 2003. With this calibration, wage inequality rises by 0.137 in CEU during the same time, compared to the .070 rise in the data (first column of table 8). These results suggest that differences in labor market policies can generates about 59% ($= (0.218 - .137)/(0.207 - 0.0699)$) of the widening in the inequality gap between CEU and US-UK during this time period.

Another dimension of the rise in wage inequality is seen in table 2 and replicated in columns 2 and 3 of the last table. The substantial part of the rise in 90-10 differential in the CEU has been at the top: the 90-50 differential is responsible for 91% of the total rise in the 90-10 differential during this period, whereas only 9% is at the lower end. A similar outcome, somewhat less extreme, is observed in the US and UK where 74% of the rise in the log 90-10 differential is due to the 90-50 differential. The model generates a similar

Table 8: Contribution of Taxes to the Rise in Wage Inequality

		Change in Log 90-10 Wage Diff.		
		Total 90-10	% 90-50	% 50-10
CEU	Data	0.0699	91%	9%
	Model	0.137	94%	6%
US/UK	Data	0.207	76%	24%
	Model	0.218	95%	5%
Difference	Data	0.137	68%	32%
	Model	0.081	97%	3%
% Explained		**59%**	**86%**	**3%**

pictures and even overstates the role of the upper end: about 94% of the rise in teh CEU and 95% in the US-UK is due to the 90-50 differential.[16]

5.1 Welfare Analysis [to be written]

6 Conclusion [to be written]

7 Appendix:

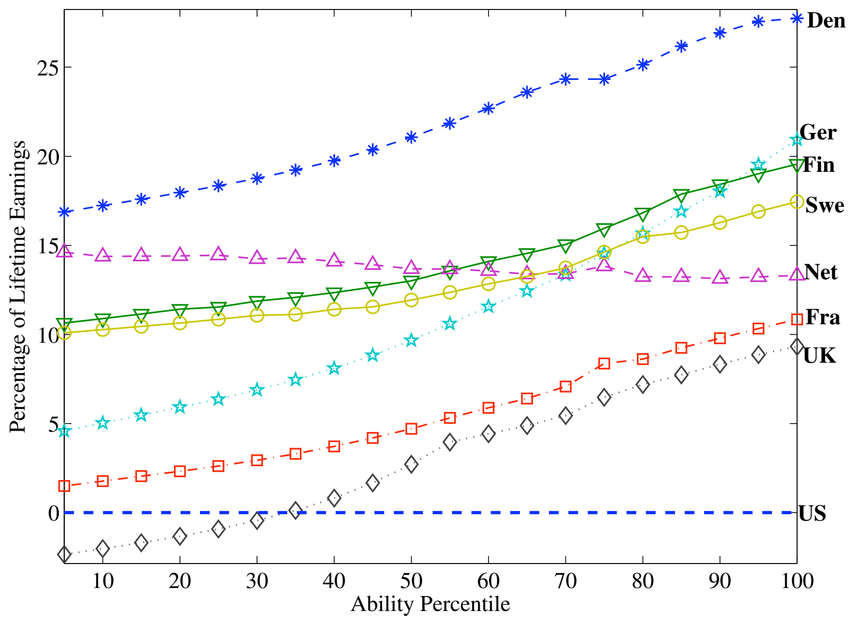
Describe details of the retirement and unemployment systems.

Definition of y^*. First,

$$y^* = [(P_L l + P_H h)(1 - i^*)]\, n^*$$

[16]The current calibration does not have idiosyncratic shocks and differences in legal minimum wages yet, which may affect the role of the lower tail in the model too.

Figure 9: Percentage of Lifetime Net Earnings Willing to Pay to Live in the US



where n^* and i^* are given by the solution to the problem below

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

References

Fatih Guvenen. An empirical investigation of labor income processes. *Review of Economic Dynamics*, 12(1):58–79, January 2009. 3

Fatih Guvenen. Learning your earning: Are labor income shocks really very persistent? *American Economic Review*, 97(3):687–712, June 2007. 3

Fatih Guvenen and Burhanettin Kuruscu. A quantitative analysis of the evolution of the u.s. wage distribution:1970-2000. NBER Working Papers 13095, National Bureau of Economic Research, Inc, May 2007. 4, 18, 24

M. Huggett, G.J. Ventura, and A. Yaron. Sources of lifetime inequality. 2007. 3

Lawrence F Katz and Kevin M Murphy. Changes in relative wages, 1963-1987: Supply and demand factors. *The Quarterly Journal of Economics*, 107(1):35–78, February 1992. 18

Figure 10: Wage Data Definitions and Sample Selection Criteria

Country	Defn	Earnings definition	Original source	Publication/data provider	Workers not covered	Note
Denmark	GHO0	Gross hourly earnings.	Tax registers (annual earnings data) and social security data data (hours worked).	The data were supplied by Professor Niels Westergård-Nielsen, Centre for Labour Economics, Aarhus Business School.	Workers with wage rates lower than 80 per cent of the minimum wage.	The data are derived from annual wage-income (including all types of taxable wage-income) recorded in tax registers, divided by actual hours worked, as recorded in a supplementary pension scheme register.
Finland	GAY0	Gross annual earnings of full-time, full-year workers.	Household survey (*Income Distribution Survey*).	Statistics Finland.	No exclusions.	
France	NAE0	Net annual earnings of full-time, full-year workers.	Salary records of enterprises. (*Déclarations Annuelles des Données Sociales*).	Institut national de la statistique et des études économiques (INSEE), *Séries longues sur les salaires*.	Agricultural and general government workers and household service workers representing 80% of salaried workers. Published results exclude apprentices, trainees, subsidised jobs and self-employed workers in unincorporated enterprises.	The data are adjusted for annual hours worked to represent full-year equivalent earnings. The data for 1981,1983,1990 are estimations by INSEE. Earnings are net of employee social security contributions but not of income tax.
Germany (Western Germany)	GMF0	Gross monthly earnings of full-time workers.	Household survey (German Socio-Economic Panel).	Secretariat calculations.	Apprentices.	Data refer to current monthly wage plus 1/12 of supplementary payments comprising 13th month pay, 14th month pay, holiday allowances and Christmas allowances.
Netherlands	GAE0	Annual earnings of full-time, full-year equivalent workers.	Enterprise survey (*Survey of Earnings*).	Sociaal-Economische Maandstatistiek, Dutch Central Bureau of Statistics.	No exclusions.	Earnings deciles are Secretariat interpolations of the published data on the distribution of employees by earnings class. Occasional payments (overtime, holiday, etc.) are included.
Sweden	GAY0	Gross annual earnings of full-year, full-time workers.	Household survey (*Income Distribution Survey*).	Statistics Sweden.	No exclusions. Employees aged 20-64 without earnings from self-employment.	New basis, consistent with 1991 change in tax base and spliced onto old series prior to 1989.
United Kingdom (Great Britain)	GWF2	Gross weekly earnings of all full-time workers (i.e. on adult or junior rates of	Enterprise survey (*New Earnings Survey*).	(former) U.K. Department of Employment.	No exclusions.	The data refer to employees whose pay was not affected by absence and include overtime and other supplementary payments.
United States	GWF0	Gross usual weekly earnings of full-time workers aged 16 and over.	Household survey (*Current Population Survey*). Unpublished BLS tabulations from the Current Population Survey and, for the published decile limits (in bold), BLS, Highlights of Women's Earnings, various editions.	U.S. Bureau of Labor Statistics.	No exclusions.	Earnings deciles are Secretariat interpolations of the unpublished BLS tabulations of the distribution of employees by earnings class.