Optimal taxation and the skill premium[*]

Konstantinos Angelopoulos
University of Glasgow

James Malley
University of Glasgow and CESifo

Apostolis Philippopoulos
Athens University of Economics & Business, CESifo

January 10, 2012

Abstract

The stylized facts suggest a negative relationship between tax progressivity and the skill premium from the early 1960s until the early 1990s, and a positive one thereafter. They also generally imply rising tax progressivity, except for the 1980s. In this paper, we ask whether optimal tax policy is consistent with these observations, taking into account the demographic and technological factors that have also affected the skill premium. To this end, we construct a dynamic general equilibrium model in which the skill premium and the progressivity of the tax system are endogenously determined, with the latter being optimally chosen by a benevolent government. We find that optimal policy delivers both a progressive tax system and model predictions which are generally consistent, except for the 1980s, with the stylized facts relating to the skill premium and progressivity. To capture the patterns in the data over the 1980s requires that we adopt a government policy which is biased towards the interests of skilled agents. Thus, in addition to demographic and technological factors, changes in the preferences of policy-makers appear to be a potentially important factor in determining the evolution of the observed skill premium.

Keywords: skill premium, optimal tax policy, government preferences
JEL Classification: E62, E65, J31

1

1 Introduction

The evolution and determinants of excess returns to skilled labor have been extensively researched over the past several decades (see e.g. Hornstein *et al.* (2005) and Acemoglu and Autor (2011) for reviews of this literature). To explain the movements in the U.S. skill premium depicted in Figure 1 below, the literature has concentrated on the importance of relative skill supply and skill-biased productivity change (see e.g. Katz and Murphy (1992), Cummins and Violante (2002), Krusell *et al.* (2000) and Autor *et al.* (2008)).[1] *Ceteris paribus*, an increase in the supply of skilled relative to unskilled labor tends to decrease the skill premium, whereas technological innovations that increase the productivity of capital raise the relative productivity of skilled labor and in turn the skill premium. To explain the relative stability and decline of the skill premium up to the early 1980s, the dominance of demographic versus technological factors is often cited, whereas the rise in the skill premium from the early 1980s is mainly attributed to the relative strength of investment specific technological change.

[Figure 1 here]

The benchmark model used in the skill premium literature allows for two types of labor, skilled and unskilled, which exhibit different degrees of complementarity with capital. One strand of recent research has focused on extending this canonical model by endogenizing skilled labour supply and technical change and another on explaining different characteristics of the distribution of wages and wage inequality over time (see e.g. Hornstein *et al.* (2005) and Acemoglu and Autor (2011) for reviews and extensions). The part played by economic policy in determining the observed evolution of the skill premium has also been considered. For instance, Cozzi and Impullitti (2010) analyze the role of fiscal policy and technical change, while Card and DiNardo (2002) and Lemieux (2006) examine the effects of minimum wage legislation. However, economic policy is not only a potential cause but may also be a result of wage inequality. This is probably most evident in policy choices relating to optimal tax design which depend on the underlying income inequality in the population (see e.g. the review papers by Mankiw

---

[1]Empirically, the skill premium is typically defined as the ratio of the wage rate of university educated workers relative to the wage rate of workers without tertiary education and the relative skill supply as the ratio of the skilled to unskilled labour force. The labour market data used in Figure 1 are from Acemoglu and Autor (2011) and the technical progress data are from Cummins and Violante (2002). More details on these data are provided later.

et al. (2009) and Diamond and Saez (2011)).[2]

A well documented feature of the US tax system is its increasing progressivity until the early 1980s, a sharp fall over the 1980s and a reversal back to rising progressivity from the 1990s (see e.g. Gouveia and Strauss (1994) and Piketty and Saez (1997)). These trends can clearly be seen in Figure 2 by examining the movement in the relative tax rates.[3] The fall in progressivity over the 1980s is usually attributed to policy reforms undertaken during the Reagan administration, in particular the Economic Recovery Tax Act of 1981.

[Figure 2 here]

Figure 2 further suggests that the relationship between the skill premium and the progressivity of the tax system is negative until the early 1990s, and positive thereafter. In this paper, we examine whether optimal tax policy can help us to capture this stylized fact in a model where both the progressivity of the tax system and the skill premium are endogenously determined. Our approach allows for a feedback effect from optimal tax policy to the skill premium via the heterogeneities in marginal propensities to save and skill-biased capital accumulation. This creates an extra channel for tax policy to redistribute income and thus increases the effectiveness of progressive taxation.

For example, when skill and returns to skill are exogenous, as typically assumed in the optimal taxation literature, the government supports the poorer by taxing the rich more, thus affecting the after-tax labor income distribution. However, when the returns to skill are endogenous, as in our model, an increase in the progressivity also results in decreased capital accumulation, since the richer households also have higher marginal propensities to save. This fall in the capital stock, in turn, increases the return to unskilled labor and thus decreases the skill premium. Therefore, in our setup, progressive taxation is more effective in supporting the income of the poor, which implies that the government can achieve its objectives with a milder progressivity.

To achieve our aims, we build on the benchmark model of skill heterogeneity, following e.g. Stokey (1996) and Krusell *et al.* (2000), where skilled

[2]The literature on optimal taxation has examined the progressivity of income taxes under different assumptions. For example, Judd (1985) shows that, in the long-run, capital income should not be taxed, whereas labor income must be positively taxed, even if capital is held by a subset of the population. This result implies that optimal taxation should be regressive with respect to income. On the other hand, skill (or ability) heterogeneity is expected to lead to optimal progressive taxes (see e.g. Conesa *et. al.* (2009)).

[3]The tax rate data used in Figure 2 are from Piketty and Saez (2007). High, middle and low refer to the average tax for households in the 80-100, 40-80 and 0-40 percentiles respectively.

labor complements capital more than unskilled labor, and extend it to allow for heterogeneity in capital holdings, given that wage inequality and wealth inequality are positively related.[4] In particular, we assume three types of agents, to reproduce the high-, middle- and low-income brackets in the tax system shown in Figure 2. We further assume that high- and middle-income earners supply skilled labor, whereas low-income earners supply unskilled labor. We allow for heterogeneity in marginal propensities to save and asset holdings by letting agents face different costs in accessing the financial markets, following the insights of the literature on credit constraints and income inequality (see e.g. Aghion and Howitt (1998)). Finally, the tax code allows the average tax rate to increase with pre-tax income, as is e.g. Atkinson and Stiglitz (1980). The progressivity, or regressivity, of the tax code is optimally chosen by a government that has access to a commitment technology and maximises the weighted average of the lifetime utility of the three agents.

When the model is calibrated to the existing skill and asset heterogeneity in the US data, we show that a Benthamite government will choose an optimal progressive tax system that is quantitatively similar to the existing US tax structure. We also establish that our model matches the predictions of the literature on the skill premium, i.e. *ceteris paribus* increases in the relative supply of skilled labor and capital augmenting technical change lead to falls and rises in the skill premium respectively. Moreover, we find that optimal tax policy reacts to such increases in the skill supply and capital-augmenting technology, by decreasing and increasing respectively the progressivity of the tax system.

The analysis of the data suggests that we can distinguish three distinct sub-periods when examining the patterns in the skill premium and the progressivity of the tax system, i.e. 1966-1979, 1980-1990 and 1991-2000. We empirically evaluate the model, under the assumption of a Benthamite government, by subjecting it to the observed joint changes in the capital augmenting technology and the relative skill supply parameters. We find that the model generally replicates the main patterns in the data regarding the skill premium and the progressivity of the tax system for the first and last sub-period.

However, to capture the patterns in the data during the 1980s requires that we depart from the assumption of a Benthamite government and instead consider partisan preferences (see e.g. Mueller (2003, ch. 19) for a review of the evidence relating to partizan effects on economic policy over

---

[4]The data suggest that the households with higher asset wealth, also earn, on average, higher labor income (see e.g. the Panel Study of Income Dynamics (PSID) data analysed in Garcia-Mila *et. al.* (2010)).

the 1980s). In particular, the stylized facts depicted in Figures 1 and 2 from 1980-1990 can be captured by assuming a bias towards the interests of skilled labor. Hence, changes in the preferences of policy-makers, in addition to the demographic and technological factors pointed out by the literature to date, appear to be an important factor in determining the evolution of the observed skill premium.

2 The Model

The economy is closed and inhabited by three heterogenous groups of households, firms and a government. The differentiated households differ with respect to their capital holdings and the type of labor services they provide. The former is driven by capital market imperfections implying that equilibrium investment will remain unequal across different groups revealing their different access to capital markets. The latter is reflected by high- and middle-income households offering skilled labor and low-income households providing unskilled labor. Moreover, skilled labor is assumed to be relatively more complementary to capital than unskilled labor. As a result, when capital rises, the skill wage premium also rises. Economic policy is chosen by a Ramsey government.

2.1 Economic agents and their roles

All households consume, work and can save in the form of capital subject to different capital transaction costs. Households also own the firms and receive profits which, due to constant returns to scale, are zero in equilibrium. Firms employ capital, skilled labor and unskilled labor to produce a single product. The government uses an income tax to finance utility-enhancing public spending, where the tax code allows for progressive or regressive taxation. All households have the same tastes so that they differ solely with respect to their economic roles or equivalently their sources of income.

2.2 Population composition

Total population size, N, is exogenous and constant. Among N, $N^h < N$ are high-income households, $N^m < N$ are middle-income households, and the rest, $N^l = N - N^h - N^m$, are low-income households, where the subscripts $h = 1, 2, ..., N^h$, $m = 1, 2, ..., N^m$ and $l = 1, 2, ..., N^l$ denote the three groups. Private firms are identical and indexed by the subscript $f = 1, 2, ..., N^f$. We assume that the number of firms equals the number of households, $N = N^f$,

or that each household owns one firm. We take this population composition as given. Finally, it is useful for what follows, to define the population shares $N^h/N \equiv n^h$, $N^m/N \equiv n^m$ and $N^l/N \equiv n^l = 1 - n^h - n^m$.

2.3 Households

The objective of each household of type $i \equiv h, m, l$ is to maximize:

$$\sum_{t=0}^{\infty} \beta^t u_{i,t} \left(C_{i,t}, z_{i,t}, \overline{G}_t \right) \tag{1}$$

where $C_{i,t}$ and $z_{i,t}$ are respectively household i's consumption and leisure; $\overline{G}_t$ is defined as average government services (i.e. total public consumption services divided by total population, N); and $0 < \beta < 1$ is the discount rate. The period utility function, $u_{i,t}(.)$, is increasing and strictly concave in all arguments. Following e.g. Conesa *et al.* (2009), the form we employ for the period utility function is:

$$u_{i,t} \left(C_{i,t}, z_{i,t}, \overline{G}_t \right) = \frac{C_{i,t}^{1-\gamma_1}}{1-\gamma_1} + \mu_1 \frac{z_{i,t}^{1-\gamma_2}}{1-\gamma_2} + \mu_2 \frac{\overline{G}_t^{1-\gamma_3}}{1-\gamma_3} \tag{2}$$

where $(\mu_1, \mu_2, \gamma_1, \gamma_2, \gamma_3) > 0$ are preference parameters. As noted by Conesa *et al.* (2009), this additively separable specification allows, consistent with the micro-evidence, relatively low labor supply elasticities to be adopted.

We define the pre-tax (capital and labor) income of each household of type $i \equiv h, m, l$ as:

$$Y_{i,t} \equiv (r_t - \delta)K_{i,t} + \pi_{i,t} + w_t^j l_{i,t} \tag{3}$$

where $K_{i,t}$ is the beginning-of-period capital held by each household type; r_t is the return to $K_{i,t}$; w_t^j is the wage rate with $j = s$ for types $i = h, m$ and $j = u$ for type $i = l$; $l_{i,t}$ and $z_{i,t} = 1 - l_{i,t}$ are respectively hours of work and leisure for each household type; $\pi_{i,t}$ is firms' dividends per household type; $0 < \delta < 1$ is the capital depreciation rate; and $K_{h,0} > K_{m,0} > K_{l,0} > 0$ are given.

We assume that each household type pays different intermediation or transaction premia due to imperfections in the capital market. Using capital market imperfections to help explain agent heterogeneity has been extensively employed by the income inequality literature (see e.g. Galor and Zeira (1993), Bénabou (1996) and Aghion and Howitt (1998, ch. 9)). In particular, we follow e.g. Persson and Tabellini (1992) and Benigno (2009) and assume a quadratic function such that a cost of $\varphi_i K_{i,t}^2$ maintains for holding physical capital $K_{i,t}$, where the parameter $\varphi_i > 0$ measures the size of

the transaction cost. We assume that higher incomes provide an insider advantage in financial transactions due, for instance, to past experience, socioeconomic background, networks, etc., leading to transaction costs, which are rank-ordered as follows $\varphi_h < \varphi_m < \varphi_l$.

Therefore, in each period, the budget constraint of each household i is:

$$C_{i,t} + K_{i,t+1} - K_{i,t} + \phi_i K_{i,t}^2 = Y_{i,t} - T(Y_{i,t}) \tag{4}$$

where the function $T(Y_{i,t})$ refers to taxes paid to the government by each household type.[5] Each household i, taking factor prices and policy variables as given, chooses $\{C_{i,t}, l_{i,t}, K_{i,t+1}\}_{t=0}^{\infty}$ to maximize (1) subject to (4) and the time constraint.

2.4 Firms

Firms use three factors of production (capital, skilled labor and unskilled labor). Each firm, f, maximizes profits:

$$\pi_{f,t} = Y_{f,t} - r_t K_{f,t} - w_t^s l_{f,t}^s - w_t^u l_{f,t}^u \tag{5}$$

where $Y_{f,t}$ is the firm's output and $K_{f,t}$, $l_{f,t}^s$ and $l_{f,t}^u$ are respectively the inputs of capital, skilled labor and unskilled labor employed by each firm f at t. Following e.g. Stokey (1996), we assume that the firm's production function is given by:

$$Y_{f,t} = A \left[\rho (A_k K_{f,t})^\nu + (1-\rho) (l_{f,t}^u)^\nu \right]^{\alpha/\nu} \left[l_{f,t}^s \right]^{1-\alpha} \tag{6}$$

where $A > 0$ and $A_k > 0$ are neutral and capital augmenting technology respectively; and $0 < \alpha, \rho < 1$ and $\nu > 0$ are technology parameters. This form captures the idea that capital is a substitute for unskilled labor, while skilled labor complements both capital and unskilled labor. When $\nu = 1$, capital and unskilled labor are perfect substitutes, whereas when $\nu = 0$, the function becomes a Cobb-Douglas for all inputs.

2.5 Government budget constraint

Assuming a balanced budget in each period, we have:

$$N\overline{G}_t = N^h T(Y_{h,t}) + N^m T(Y_{m,t}) + N^l T(Y_{l,t}). \tag{7}$$

---

[5]Note that the tax function is defined below when we present the government sector.

Following, e.g. Atkinson and Stiglitz (1980), the income tax function, $T(Y_{i,t})$, combines a flat marginal tax rate with a transfer to all individuals:

$$T(Y_{i,t}) = \tau_t Y_{i,t} - R_t \tag{8}$$

where τ_t and R_t are policy instruments.[6] That is, $0 < \tau_t < 1$ is the marginal tax rate and R_t (when positive) is the minimum income guaranteed by the government. This function allows for progressive taxation in the sense that the average tax rate increases with pre-tax income. Note that for the government budget constraint to be satisfied only two of the three policy instruments, τ_t, R_t, and $\overline{G}_t$ can be set independently.

2.6 Decentralized Competitive Equilibrium (DCE)

In the DCE, all households maximize welfare, all firms maximize profits, the markets for capital, skilled and unskilled labor, dividends and goods clear and all constraints are satisfied. This is for any feasible policy and initial conditions for the state variables. The ten optimality conditions summarizing the DCE are presented in the Appendix.

2.7 Ramsey policy and equilibrium

The government chooses policy instruments, $\{\tau_t, R_t, \overline{G}_t\}_{t=0}^{\infty}$, to maximize a weighted average of the three agents' welfare functions:

$$\omega_h \sum_{t=0}^{\infty} \beta^t u_h \left(C_{h,t}, l_{h,t}, \overline{G}_t \right) + \omega_m \sum_{t=0}^{\infty} \beta^t u_m \left(C_{m,t}, l_{m,t}, \overline{G}_t \right) + \tag{9}$$

$$+ (1 - \omega_h - \omega_m) \sum_{t=0}^{\infty} \beta^t u_l \left(C_{l,t}, l_{l,t}, \overline{G}_t \right)$$

subject to the DCE conditions presented in the Appendix. The fixed parameters $0 \le \omega_h, \omega_m \le 1$ are the utility preference weights. In the Benthamite case, $\omega_h = n^h$, $\omega_m = n^m$ and $(1 - \omega_h - \omega_m) = n^l$.

3 Stylized facts, calibration and solution

In this section, we further examine the data regarding the skill premium and the progressivity of the tax system. This discussion informs the calibration

---

[6]Imposing a flat tax with a universal lump-sum transfer is clearly a restriction on tax policy and is used here for simplicity (see e.g. Diamond and Saez (2011) for a discussion of the "optimal" tax formula). However, it allows us to capture, in a model with a stylized skill heterogeneity, the overall progressivity of the tax system and its relationship with the skill premium.

which aims to reproduce the main stylized facts over the entire sample period, 1966-2000.[7] It also helps to contexualize the comparative static analysis of the next section which, under different assumptions regarding demography, technology and political preferences, aims to match the key empirical observations over 1966-1979, 1980-1990 and 1991-2000.

3.1 Stylized facts

In Table 1 we summarize the main empirical observations regarding tax rates, the skill premium, relative skill supply and capital augmenting technical change. Referring to the tax rates data in Table 1, the first point to note is that the progressivity of the US tax system is reflected by the relationship between the effective average tax rates applying to the three income groups considered, i.e. $\tau^h > \tau^m > \tau^{l}$[8]. A convenient way to measure the extent of the progressivity of the tax system is to look at the ratios of the tax rates that apply to the different income groups, which are denoted as τ^h/τ^m, τ^h/τ^l and τ^m/τ^l. We also report the ratio of the effective tax rate that applies to skilled labor relative to unskilled, as $\tau^{h,m}/\tau^l$.[9] In general, an increase in these ratios captures an increase in the progressivity of the tax system, with respect to the income distribution that we consider.

The averages of the tax rates and their ratios over each period suggest that the progressivity of the tax system has generally increased and more markedly so in the 1990s. However, the average levels over the sub-periods mask different growth trends within these periods, which becomes apparent when we examine Figure 2. In particular, progressivity was increasing over 1966-1979, decreasing over 1980-1990 and then again increasing over 1991-2000. This can be seen by referring to the average growth rates of the tax ratios, denoted as $g\left(\frac{\tau^h}{\tau^m}\right)$, $g\left(\frac{\tau^h}{\tau^l}\right)$, $g\left(\frac{\tau^m}{\tau^l}\right)$ and $g\left(\frac{\tau^{h,m}}{\tau^l}\right)$ in Table 1. These ratios have generally been increasing over the periods considered, with the notable exception of the 1980s, where all measures suggest a declining progressivity. As pointed out in the Introduction, this fall in progressivity has been attributed to policy reforms undertaken during the Economic Recovery

---

[7]Note that the start date is restricted by the tax data and the end date by the technological change data referred to in Figures 1 and 2.

[8]As noted earlier, the data used to construct the measures reported in Table 1 are obtained from Piketty and Saez (2007). The τ^h rate is the effective average tax rate that applies to the top quantile, τ^m has been calculated as the average tax rate of the second and third quantiles, while τ^l has been calculated as the average tax rate of the fourth and fifth quantiles.

[9]The effective tax rate that applies to skilled labour $\left(\tau^{h,m}\right)$ is calculated as the weighted average of the τ^h and τ^m tax rates.

Tax Act of 1981 (see, e.g. the discussion in Gouveia and Strauss (1994).[10]

Table 1: Historical Data

	1966-2000	1966-1979	1980-1990	1991-2000
Levels and relative tax rates				
τ^h	0.2914	0.2760	0.2925	0.3116
τ^m	0.2175	0.2085	0.2249	0.2218
τ^l	0.1524	0.1614	0.1544	0.1376
τ^h/τ^m	1.3401	1.3249	1.3005	1.4050
τ^h/τ^l	1.9347	1.7148	1.8980	2.2828
τ^m/τ^l	1.4397	1.2951	1.4588	1.6209
$\tau^{h,m}/\tau^l$	1.6047	1.4350	1.6052	1.8415
Skill premium and relative skill supply				
w^s/w^u	1.6336	1.5444	1.6003	1.8076
$(n_h + n_m)/n_l$	1.5547	1.4251	1.6024	1.6987
Growth rates (g) and correlations (ρ)				
$g\left(\frac{\tau^h}{\tau^m}\right)$	0.0009	-0.0018	-0.0054	0.0109
$g\left(\frac{\tau^h}{\tau^l}\right)$	0.0124	0.0161	-0.0134	0.0267
$g\left(\frac{\tau^m}{\tau^l}\right)$	0.0115	0.0180	-0.0081	0.0157
$g\left(\frac{\tau^{h,m}}{\tau^l}\right)$	0.0119	0.0172	-0.0102	0.0201
$g\left(\frac{w^s}{w^u}\right)$	0.0062	-0.0014	0.0118	0.0074
$g\left(\frac{n_h+n_m}{n_l}\right)$	0.0073	0.0093	0.0059	0.0046
$g\left(A_k\right)$	0.0309	0.0150	0.0318	0.0430
$\rho\left(\frac{w^s}{w^u},\frac{\tau^h}{\tau^m}\right)$	0.5682	-0.1155	-0.7076	0.8385
$\rho\left(\frac{w^s}{w^u},\frac{\tau^h}{\tau^l}\right)$	0.6917	-0.5815	-0.9344	0.7963
$\rho\left(\frac{w^s}{w^u},\frac{\tau^m}{\tau^l}\right)$	0.6399	-0.5332	-0.8911	0.7348
$\rho\left(\frac{w^s}{w^u},\frac{\tau^{h,m}}{\tau^l}\right)$	0.6696	-0.5633	-0.9235	0.7685

The next observation in Table 1 relates to the evolution of the skill premium and its main determinants, i.e. capital augmenting technical change and the relative supply of skilled workers, which have also been plotted in Figure 1. The technical change data, obtained from Cummins and Violante (2002), denoted by $g\left(A_k\right)$ for the average growth rate, suggest that capital-augmenting technology has been increasing over the whole period at an in-

---

[10]Note that the income distribution that we consider (i.e. top 20%, middle 40% and low 40%) blurs the evolution of the top income tax rates that apply to the highest percentiles (e.g. the top 1%). As Piketty and Saez (2007) show, there has been a consistent reduction in these tax rates over the whole time period that we consider.

creasing rate.[11] The skill supply and skill premium data are from Acemoglu and Autor (2011) and have been extensively analysed in e.g. Autor *et al.* (2008) and Acemoglu and Autor (2011).[12] As evidenced by $g\left(\frac{n_h+n_m}{n_l}\right)$ over the three sub-periods considered, Table 1 shows that the relative skill supply has been increasing over the whole period, but at a decreasing rate. The skill premium, w^s/w^u, on the other hand, decreased over the first sub-period, while it increased after 1980, see e.g. $g\left(\frac{w^s}{w^u}\right)$ in Table 1.

The above changes in the progressivity of the tax system and the skill premium suggest that there should be a negative relationship between these two series over the first two sub-periods and a positive relationship in the third. This is indeed confirmed when we look at the correlations between the skill premium and the ratios of the tax rates, shown in the last four rows of Table 1. It is useful to note, however, that the underlying relationship is not the same over the first two sub-periods, despite the negative correlation in both cases. In particular, over 1966-1979, progressivity has been increasing, while the skill premium has been decreasing. On the other hand, over 1980-1990, progressivity has been decreasing, while the skill premium has been increasing. Therefore, these two sub-periods differ fundamentally in the relationship between progressivity and the skill premium, which motivates us to examine them separately in the next section.

3.2 Calibration

The model parameters, reported in Table 2, are calibrated so that the model generally captures the data averages over the whole period, 1966-2000, reported in Table 1 above and Table 3 below. The population shares are chosen so that the three income groups in the model correspond to the high-, middle- and low-income groups shown in the data. Therefore, we set n_h, n_m and n_l to 0.2, 0.4 and 0.4 respectively and assume that the high- and middle-income groups supply skilled labor and own a larger share of aggregate capital. The first assumption implies a relative skill supply of 1.5, which is consistent with the data average reported in Table 1. To capture the wealth distribution in the data, we calibrate the transaction cost parameters, so that $\phi_m = 2.5 \times \phi_h$

---

[11]To construct this measure, we follow He (2011) and use the quality-adjusted price index of total investment (in equipment and structures) and the official NIPA price index of consumption (nondurable and services) from Cummins and Violante (2002). The resulting capital augmenting technology index was also re-intialized at 1966.

[12]Following Autor *et. al.* (2008) and Acemoglu and Autor (2011), the relative skill supply is calculated as the ratio of college equivalent to non-college equivalent labour supply in efficiency units. The skill premium is defined as the ratio of college to high school wages. Note that in Figure 1 we plot the natural logarithms of these ratios.

and $\phi_l = 5 \times \phi_h$, which implies that high-income households own two and a half times as much capital stock as the middle-income households and five times as much as the low-income households. This again coheres with PSID data from the 80s (see e.g. Garcia-Milà *et al.* (2010)) and US Census data from the 90s (see e.g. Eller and Fraser (1995)). Implicitly, our model and calibration assume that more skilled households also own more assets, which is, on average, similar to PSID data on average wages for different wealth groups (see e.g. Table 2 in Garcia-Milà *et al.* (2010)). The size of transaction costs, as captured by ϕ_h, is calibrated so that, given a commonly used depreciation parameter and rate of time preference, it implies a steady-state total capital to output ratio that is consistent with the data.[13]

The levels of total factor productivity and capital augmenting technology are both normalised to unity. The parameters ν, α and ρ are important in determining the skill premium, as they affect the returns to capital and the two types of labor, as well as the complementarity between the three factors. We follow Stokey (1996) and set ν and ρ to 0.5 and 0.4 respectively. The latter implies an elasticity of substitution between capital and unskilled labor, $1/(1-\nu)$, of 2, which is well within the range of estimated elasticities of substitution reported in the literature (see e.g. Stokey (1996) and Hornstein *et al.* (2005) for reviews of these studies). The remaining parameter in the production function, α, is set to 0.5, to match the average level of skill premium in the data.

Table 2: Parameters

A	A_k	α	β	γ_1	γ_2	γ_3	δ	μ_1
1.000	1.000	0.500	0.972	2.500	2.500	2.500	0.083	1.920
μ_2	ν	ρ	ϕ_h	ϕ_m	ϕ_l	n_h	n_m	n_l
0.025	0.500	0.400	0.0015	$\phi_h \times 2.5$	$\phi_h \times 5$	0.200	0.400	0.400

The weight on leisure relative to private consumption, μ_1, is as in Conesa *et al.* (2009). We calibrate the weight given to public relative to private consumption, μ_2, to obtain a public consumption spending to output ratio of about 16%. Finally, the curvature parameters, $\gamma_1, \gamma_2, \gamma_3$, in the utility function are set in the middle of the commonly employed range, i.e. $(2-3)$.

3.3 Steady-state

Assuming Benthamite policy, the steady-state solution of the model is detailed in Table 3. We report the solution for the two policy instruments, τ and R, together with the implied effective average tax rates for each type of

---

[13]The values of $\delta = 8.33\%$ and $\beta = 0.972$ are as in Conesa *et. al.* (2009).

agent and the resulting tax ratios that measure the progressivity of the tax system along with data averages for the available measures.

Table 3: Steady-State

	Data	Model
Tax rates & skill premium		
τ	n.a.	0.3530
R	n.a.	0.0453
τ^h	0.2914	0.2570
τ^m	0.2175	0.2298
τ^l	0.1524	0.1622
τ^h/τ^m	1.3401	1.1183
τ^h/τ^l	1.9347	1.5845
τ^m/τ^l	1.4397	1.4169
$\tau^{h,m}/\tau^l$	1.6047	1.3174
w^s/w^u	1.6336	1.6154
Asset distribution & Frish elasticities		
K_h/K_m	2.2180	2.5000
K_h/K_l	4.7188	5.0000
K_m/K_l	2.1278	2.0000
f_h	n.a.	0.4311
f_m	n.a.	0.3790
f_l	n.a.	0.3019

The first result that can be seen in Table 3 is that the assumed heterogeneities imply an optimal progressive tax system, which has qualitative and quantitative features that are similar to the data averages. This is an important prediction of the model. As discussed earlier, under skill heterogeneity, optimal tax policy has already been shown to be progressive (see e.g. Conesa *et al.* (2009) and Diamond and Saez (2011) for a review). However, here we also allow the government to optimally affect the skill premium and thus pre-tax wage inequality. Therefore, optimal progressive taxation is obtained for a Benthamite government that endogenizes the feedback effect of its policies on the returns to skill and income distribution.

The feedback effect creates an extra channel for tax policy to redistribute income and thus increases the effectiveness of progressive taxation. In particular, when skill and returns to skill are exogenous, as typically assumed in the optimal taxation literature, the government supports the poorer by taxing the rich more, thus affecting the after-tax labor income distribution. However, when the returns to skill are endogenous, as in our model, an increase in the progressivity also results in decreased capital accumulation, since the

richer households also have higher marginal propensities to save. This fall in the capital stock, in turn, increases the return for unskilled labor and thus decreases the skill premium. Therefore, in our setup, progressive taxation is more effective in supporting the income of the poor, which implies that the government can achieve its objectives with a milder progressivity.

The level of the effective tax rates predicted by the model is similar to that in the data for the middle and low income groups, but is lower, compared to the data for the top income group, which also implies that the predicted progressivity is somewhat smaller than in the data. However, if the very high tax rates that apply to the top 1% of the distribution are excluded from the calculation of τ^h, the data averages for all tax rates are very similar to the model predictions (see e.g. Piketty and Saez (2007)).

Note also that the implications of the modeling assumptions and calibration cohere with the average skill premium and asset heterogeneity present in the data. Moreover, consistent with the findings from the microeconometric literature (see e.g. Browning *et al.* (1999)), the model produces Frish labor supply elasticities that are less than one.

4 Optimal progressivity and the skill premium

4.1 Model predictions

We first examine the predictions of the model relating to the skill premium and optimal tax policy with respect to exogenous changes in the relative skill supply, capital augmenting technology and government preferences. As pointed out in the introduction, the literature suggests that the first two factors are important determinants of the skill premium. However, the response of the optimal tax policy to these changes has not been examined previously. Moreover, although political preferences have been linked to the tax system in the political economy literature (see e.g. Drazen (2000), Persson and Tabellini (2000), Mueller (2003) and Winer and Hettich (2003) for reviews of this literature), their effect on the skill premium via the choice of the progressivity of the tax system has also not been systematically studied.

The model predictions for the steady-state solution with respect to the changes in the relevant parameters are presented in Table 4. We first present the solution for the base parameterization explained above, followed by the results obtained by a 1% rise in the relative skill supply, the capital augmenting technology and the weight for skilled households in the government's objective function.[14] To facilitate the discussion of the results, we also present,

---

[14]Note that results obtained from decreases in the above factors are symmetric and thus

14

for each case, the percent deviation from the base solution for the skill premium and the progressivity measures.

Table 4: Model Predictions

	Base	1% rise in: $(n_h + n_m)/n_l$	1% rise in: A_k	1% rise in: $\omega_h + \omega_m$
Tax rates & skill premium				
τ	0.3530	0.3486	0.3557	0.3489
R	0.0453	0.0442	0.0461	0.0442
τ^h	0.2570	0.2551	0.2579	0.2554
τ^m	0.2298	0.2285	0.2303	0.2289
τ^l	0.1622	0.1634	0.1611	0.1629
τ^h/τ^m	1.1183	1.1164	1.1202	1.1157
τ^h/τ^l	1.5845	1.5608	1.6007	1.5675
τ^m/τ^l	1.4169	1.3980	1.4290	1.4049
$\tau^{h,m}/\tau^l$	1.3174	1.3042	1.3261	1.3085
w^s/w^u	1.6154	1.6055	1.6187	1.6187
Asset distribution & Frish elasticities				
K_h/Y	0.8972	0.9038	0.8989	0.8998
K_m/Y	0.7177	0.7230	0.7191	0.7198
K_l/Y	0.3589	0.3579	0.3595	0.3599
f_h	0.4311	0.4305	0.4332	0.4305
f_m	0.3790	0.3782	0.3806	0.3784
f_l	0.3019	0.3012	0.3031	0.3006
Progressivity & skill premium relative to base				
$\widehat{d}\left(\tau^h/\tau^m\right)$	0.0000	-0.0017	0.0017	-0.0023
$\widehat{d}\left(\tau^h/\tau^l\right)$	0.0000	-0.0150	0.0102	-0.0108
$\widehat{d}\left(\tau^m/\tau^l\right)$	0.0000	-0.0133	0.0085	-0.0085
$\widehat{d}\left(\tau^{h,m}/\tau^l\right)$	0.0000	-0.0100	0.0066	-0.0067
$\widehat{d}\left(w^s/w^u\right)$	0.0000	-0.0061	0.0020	0.0020

4.1.1 Demographic changes

Other things equal, an increase in the relative supply of skilled labor is expected to lead to a fall in the skill premium, which is indeed confirmed in Table 4. These results also show that a Benthamite government finds it optimal to decrease the progressivity of the tax system. The fall in the skill premium implies that wage inequality falls and thus the incentive for the government to redistribute income through tax policy is weaker. Moreover, given

---

not presented here to save on space.

that a larger proportion of the population is now skilled, the Benthamite government places relatively greater weight on the welfare of the skilled income groups. Therefore, both factors contribute to a less progressive tax system.

4.1.2 Technology changes

Ceteris paribus, an improvement in capital augmenting technology, since it is skill-biased, is expected to increase the skill premium. This is reflected in Table 4 which also shows that the optimal response of policy is to increase the progressivity of the tax system. A rise in the skill premium implies an increase in wage inequality. Thus a Benthamite government will find it optimal to increase the progressivity of taxation to redress this negative income effect on the unskilled.

It is noteworthy that in both comparative static exercises considered thus far, the progressivity of the tax system and the skill premium should move in the same direction. That is, other things equal, the model with a Benthamite government suggests that there should be a positive relationship between the progressivity and the skill premium.

4.1.3 Government preference changes

Governments need not have Benthamite preferences. For instance, if governments have partisan incentives, they might choose to support one population group more than another (see e.g. Drazen (2000) Persson and Tabellini (2000), Mueller (2003) and Winer and Hettich (2003) for reviews of this literature)). Other things equal, an increase in the weight for skilled households in the government's objective function is expected to lead to a fall in the progressivity of the tax system which is again confirmed in Table 4. Moreover, as the results in Table 4 make clear, this policy change also has effects on the skill premium via the feedback effect of the progressivity of the tax system. In particular, given the decline in the progressivity of the income tax, capital accumulation has increased. In turn, this benefits skilled labor more than unskilled, given the higher complementarity of the former with capital. In other words, falls in the progressivity of the tax system, driven by preferential treatment of high income groups by the government, will be skill-biased and thus increase the skill premium. In this sense, their effect on the skill premium is similar to the effect of capital augmenting technology. However, it is interesting to note that such political changes result in a negative relationship between the skill premium and progressivity, since increases in progressivity result in decreases in the skill premium and vice-versa.

4.2 Political preferences, taxation & the skill premium

Having established the model predictions and the main properties in the data regarding the relationship between the skill premium and the progressivity of the tax system, we now turn to evaluate whether the patterns in the data cohere with Benthamite policy, or, whether there is evidence that political preferences have contributed to the skill premium.[15] The analysis of the data in the previous section suggests that we can distinguish three sub-periods when examining the patterns in the skill premium and the progressivity of the tax system, 1966-1979, 1980-1990 and 1991-2000. We therefore subject the base calibration to joint changes in the capital augmenting technology and the relative skill supply parameters that match the average change in the data for each of these sub-periods.

The results from such counter-factual experiments are presented in Table 5, where, for ease of comparison, we again present the base calibration results in the first column and denote the percent deviation between the base and the new steady-state solution by $\widehat{d}(.)$.[16] These findings suggest that the general patterns in the skill premium and progressivity in the data are consistent with optimal, Benthamite policy over 1966-1979 and 1991-2000, but not over the middle period, 1980-1990.

More specifically, using actual changes in capital augmenting technology and relative skill supply for the 1966-1979 period, the model predicts a decline in the skill premium and a rise in the progressivity of the tax system as found in the data. Moreover, we next find that the model predictions for the 1991-2000 period are broadly consistent with the data, i.e. both the skill premium and the progressivity of the tax system rise. For both the 1966-1979 and 1991-2000 periods, the predicted decline/rise in the skill premium is quantitatively similar to what is observed in the data as well. Finally, for the 1991-2000 period, the predicted increase in $\tau^{h,m}/\tau^l$ is close to the data, but the model

---

[15]Here, we take asset and skill heterogeneity, as well as capital augmenting technology, as given and thus unaffected by government's policy. In general, government policy can affect both capital augmenting technology and the relative skill supply (see e.g. Cozzi and Impulliti (2010) and He and Liu (2008), respectively). In our empirical evaluation of the model, we control for the quantitative importance of these forcing variables by calibrating the model to the observed changes in both.

[16]Note that the data presented in Table 1 referred to the average growth rate of the skill premium and the measures of the progressivity of the tax system over each sub-period. To make our results comparable to those averages, we obtain the new steady-state solutions reported in Table 5, given a change in the parameters that capture the average growth in relative skill supply and capital augmenting technology. We then calculate the percent deviation between the base and the new steady-state solution for the skill premium and the tax progressivity measures.

17

under-predicts the change in the progressivity in the 1966-979 period.[17]

In contrast to the above finding, in the 1980-1990 period, the model predicts a rise in both the skill premium and the progressivity of the tax system, contrary to the negative relationship between these variables recorded in the data. In particular, the rise in the skill premium, driven by the quantitatively dominant increase in capital augmenting technology, is significantly lower in the model than in the data. On the other hand, we observe a decline in progressivity over the 1980s, which is inconsistent with a Benthamite government.

The observed pattern of a declining progressivity in a period of a rising skill premium does cohere, however, with optimal policy chosen by a government that has a partisan bias in favor of the skilled groups of the population. More specifically, if we add to the changes in capital augmenting technology and relative skill supply, a rise in the relative weight to the skilled income groups in the objective function of the government, the model can generate optimal progressivity and skill premium patterns that are seen in the data. This is shown in the last column in Table 5, where in addition to period specific changes in relative skill supply and capital augmenting technology, an increase in the weight to the skill groups of 3.5% is applied. The latter captures a bias of the Republican governments over the eighties in favor of higher income (or skilled) groups that is sufficient to generate a decline in $\tau^{h,m}/\tau^l$ by 1%.[18] As can be seen, under such partisan preferences, the progressivity of the tax system declines in accordance with the data and this allows the model to match the negative correlation between progressivity and the skill premium. Moreover, the increase in the skill premium is even higher under the skill-biased policy, due to the progressivity channel discussed in the previous sub-section. As a result, the change in the skill premium is also

---

[17]Our setup cannot account for the average decline in τ^h/τ^m in this period, but this seems to be an outlier in the data, driven by the very low computed effective tax rates for the middle income groups at the beginning of the period. Given that the main heterogeneity here derives from the skilled/unskilled decomposition, we focus our analysis in this section on the relevant measure of progressivity, $\tau^{h,m}/\tau^l$.

[18]A partisan bias in economic policy in favor of higher income groups over the 1980s has also been documented in the literature (see e.g. Mueller (2003, ch. 19)). However, using this literature, it is difficult to quantify this bias in a way that is consistent with our model. Therefore, in the last column in Table 5 we demonstrate the potential of the model to match the stylized facts, assuming a partisan bias sufficiently high to capture the average observed decline in $\tau^{h,m}/\tau^l$ by 1% over the 1980s.

quantitatively consistent with the data.

Table 5 : Government Preferences

		Benthamite			Skill-biased
	Base	1966-1979	1980-1990	1991-2000	1980-1990
Tax rates & skill premium					
τ	0.3530	0.3530	0.3589	0.3623	0.3442
R	0.0453	0.0455	0.0472	0.0483	0.0433
τ^h	0.2570	0.2567	0.2588	0.2601	0.2532
τ^m	0.2298	0.2293	0.2304	0.2311	0.2273
τ^l	0.1622	0.1618	0.1596	0.1582	0.1624
τ^h/τ^m	1.1183	1.1195	1.1233	1.1256	1.1139
τ^h/τ^l	1.5845	1.5864	1.6220	1.6439	1.5592
τ^m/τ^l	1.4169	1.4171	1.4440	1.4605	1.3997
$\tau^{h,m}/\tau^l$	1.3174	1.3179	1.3371	1.3489	1.3044
w^s/w^u	1.6154	1.6112	1.6200	1.6249	1.6322
Asset distribution & Frish elasticities					
K_h/Y	0.8972	0.9058	0.9063	0.9072	0.9156
K_m/Y	0.7177	0.7247	0.7251	0.7258	0.7325
K_l/Y	0.3589	0.3590	0.3604	0.3612	0.3641
f_h	0.4311	0.4336	0.4373	0.4397	0.4352
f_m	0.3790	0.3807	0.3837	0.3855	0.3817
f_l	0.3019	0.3031	0.3053	0.3067	0.3003
Progressivity & skill premium relative to base					
$\widehat{d}\left(\tau^h/\tau^m\right)$	0.0000	0.0010	0.0044	0.0065	-0.0039
$\widehat{d}\left(\tau^h/\tau^l\right)$	0.0000	0.0012	0.0237	0.0375	-0.0160
$\widehat{d}\left(\tau^m/\tau^l\right)$	0.0000	0.0001	0.0191	0.0308	-0.0121
$\widehat{d}\left(\tau^{h,m}/\tau^l\right)$	0.0000	0.0004	0.0150	0.0239	-0.0098
$\widehat{d}\left(w^s/w^u\right)$	0.0000	-0.0026	0.0029	0.0059	0.0103

   To summarise, by endogenizing tax policy, we find evidence that, in addition to relative skill supply and capital augmenting technology, political preferences have affected the skill premium. This is particularly evident in accounting for the patterns of the skill premium and progressivity in the 1980s, a period where room for non-market or policy influences are widely believed to have had an effect on the skill premium (see e.g. Autor *et al.* (2008)). At the same time, it is important to note that the political preferences relating to the progressivity of the tax system in our setup do not substitute for the market forces affecting wage inequality. Instead, they work through the supply and demand for skilled and unskilled labour by affecting their marginal productivities differentially.

5 Conclusions

In this paper we have shown that a Benthamite government would choose a progressive tax system when the main characteristics of asset and skill inequality match the stylized facts of the US economy over the past 40 years. By allowing the skill premium to be endogenously determined along with tax policy, we have also demonstrated that the main characteristics of the progressivity of income taxation in the US are consistent with optimal policy chosen by a government that cares equally about all citizens. Moreover, our results suggest that the progressivity of the optimally chosen tax system has contributed to pre-tax wage inequality in addition to changes in skill supply and capital-augmenting technology.

However, the assumption of a Benthamite government is not supported by the data for the whole sample. In particular, the 1980s is consistent with a tax policy that placed relatively greater weight on the welfare of the skilled segments of the labor force. Therefore, to capture the stylized facts in this period requires that we also allow for political economy considerations in the form of partizan preferences on the part of the policymaker.

An obvious limitation of the current paper is that we have taken government preferences as given. Thus, to shed more light into the political mechanism generating the observed politico-economic equilibrium, a useful extension would be to examine the joint determination of the government's preferences along with wage inequality and the skill premium. This might be achieved in a model incorporating incentives for political parties to support their constituents as in the literature initiated by Persson and Tabellini (1992, 1994) and Alesina and Rosenthal (1995).

References

[1] Acemoglu, D. and D. Autor (2011). 'Skills, tasks and technologies: implications for employment and earnings', in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, 4b, 1043-1171, Amsterdam: Elsevier.

[2] Aghion, P. and P. Howitt (1998). *Endogenous Growth Theory*, Cambridge, MA: MIT Press.

[3] Alesina, A. and H. Rosenthal (1995). *Partisan Politics, Divided Government, and the Economy*, Cambridge, Cambridge University Press.

[4] Atkinson, A. and J. Stiglitz (1980). *Lectures on Public Economics*, McGraw Hill, New York.

[5] Autor, D., Katz, L. and M. Kearney (2008). 'Trends in U.S. Wage Inequality: Revising the Revisionists', *The Review of Economics and Statistics*, 90, 300-323.

[6] Bénabou, R. (1996). 'Inequality and growth', in B. Bernanke and J. Rotemberg (eds.), *NBER Macroeconomics Annual 1996,* 11-92, Cambridge, MA: MIT Press.

[7] Benigno, P. (2009). 'Price stability with imperfect financial integration', *Journal of Money, Credit and Banking*, 41, 121-149.

[8] Browning, M., Hansen, L. and J. Heckman (1999). 'Micro data and general equilibrium models', in J. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, 1, 543-633, Amsterdam: Elsevier.

[9] Card, D. and J. DiNardo (2002). 'Skill-biased technological change and rising wage inequality: some problems and puzzles', *Journal of Labor Economics*, 20, 733–783.

[10] Conesa, J., Kitao, S. and D. Krueger (2009). 'Taxing capital? Not a bad idea after all!', *American Economic Review*, 99, 25-48.

[11] Cozzi, G. and G. Impullitti (2010). 'Government spending composition, technical change, and wage inequality', *Journal of European Economic Association*, 8, 1325-1358.

[12] Diamond, P. and E. Saez (2011). 'The case for a progressive tax: from basic research to policy recommendations', *Journal of Economic Perspectives*, 25, 165–190.

[13] Drazen, A. (2000). *Political Economy in Macroeconomics.* Princeton University Press, Princeton, New Jersey.

[14] Cummins, J. and G. Violante (2002). 'Investment-specific technical change in the US (1947-2000): measurement and macroeconomic consequences', *Review of Economic Dynamics*, 5, 243-284.

[15] Eller, T. and W. Fraser (1995). *Asset Ownership of Households: 1993,* U.S. Bureau of the Census. Current Population Reports, P70-47, U.S. Government Printing Office, Washington, DC.

[16] Galor, O. and J. Zeira (1993). 'Income distribution and macroeconomics', *Review of Economic Studies*, 60, 35-52.

[17] Garcia-Milà, T., Marcet, A. and E. Ventura (2010). 'Supply side interventions and redistribution', *Economic Journal*, 120, 105-130.

[18] Gouveia, M. and R. Strauss (1994). 'Effective federal individual income tax functions: an explanatory empirical analysis', *National Tax Journal*, 47, 317-339.

[19] He, H. (2011). 'What drives the skill premium: Technological change or demographic variation?', Working Paper 200911, University of Hawaii at Manoa.

[20] He, H. and Z. Liu (2008). 'Investment specific technological change, skill accumulation, and wage inequality, *Review of Economic Dynamics*, 11, 314-334.

[21] Hornstein, A., Krusell, P. and G. Violante (2005). 'The effects of technical change on labor market inequalities', in P. Aghion and S. Durlauf (eds.), *Handbook of Economic Growth*, 1, 1275-1370. Amsterdam: Elsevier.

[22] Judd, K. (1985). 'Redistributive taxation in a simple perfect foresight model', *Journal of Public Economics*, 28, 59-83.

[23] Katz, L. and K. Murphy (1992). 'Changes in relative wages, 1963-1987: supply and demand factors', *The Quarterly Journal of Economics*,107, 35-78.

[24] Krusell, P., Ohanian, L., Ríos-Rull J. and G. Violante (2000). 'Capital–skill complementarity and inequality: a macroeconomic analysis', *Econometrica*, 68, 1029–1053.

[25] Lemieux, T. (2006). 'Increasing residual wage inequality: composition effects, noisy data, or rising demand for skill?,' *American Economic Review*, 96, 461-498.

[26] Mankiw, G., Weinzierl, M. and D. Yagan (2009). 'Optimal taxation in theory and practice', *Journal of Economic Perspectives*, 23, 147–74.

[27] Mueller, D. (2003). *Public Choice III*. Cambridge: Cambridge University Press.

[28] Piketty, T. and E. Saez (2007). 'How progressive is the U.S. federal tax system? A historical and international perspective', *Journal of Economic Perspectives*, 21, 3-24.

[29] Persson, T. and G. Tabellini (1992). 'The politics of 1992: fiscal policy and European integration', *Review of Economic Studies*, 59, 689–701.

[30] Persson, T. and G. Tabellini (1994). 'Representative democracy and capital taxation', *Journal of Public Economics*, 55, 53-70.

[31] Persson, T. and G. Tabellini (2000). *Political Economics: Explaining Economic Policy*. MIT Press. Cambridge

[32] Stokey, N. (1996). 'Free trade, factor returns, and factor accumulation', *Journal of Economic Growth*, 1, 421-447.

[33] Winer, S. and W. Hettich (2003). 'The political economy of taxation: positive and normative analysis when collective choice matters', in C. W. Rowley and F. Schneider (eds.), *The Encyclopedia of Public Choice*, Dordrecht, Netherlands: Kluwer Academic Publishers.

6   Appendix

The DCE is summarized by the following ten equations:

$$\frac{1}{(C_{h,t})^{\gamma_1}} = \beta \frac{1}{(C_{h,t+1})^{\gamma_1}} \left[1 - 2\phi_h K_{h,t+1} + (1 - \tau_{t+1})(r_{t+1} - \delta)\right] \qquad (A1)$$

$$\frac{\mu_2}{(1 - l_{h,t})^{\gamma_2}} = \frac{\mu_1}{(C_{h,t})^{\gamma_1}} (1 - \tau_t) w_t^s \qquad (A2)$$

$$\frac{1}{(C_{m,t})^{\gamma_1}} = \beta \frac{1}{(C_{m,t+1})^{\gamma_1}} \left[1 - 2\phi_m K_{m,t+1} + (1 - \tau_{t+1})(r_{t+1} - \delta)\right] \qquad (A3)$$

$$\frac{\mu_2}{(1 - l_{m,t})^{\gamma_2}} = \frac{\mu_1}{(C_{m,t})^{\gamma_1}} (1 - \tau_t) w_t^s \qquad (A4)$$

$$C_{m,t} + K_{m,t+1} - K_{m,t} + \phi_m K_{m,t}^2 = (1 - \tau_t)[(r_t - \delta)K_{m,t} + w_t^s l_{m,t}] + R_t \qquad (A5)$$

$$\frac{1}{(C_{l,t})^{\gamma_1}} = \beta \frac{1}{(C_{l,t+1})^{\gamma_1}} \left[1 - 2\phi_l K_{l,t+1} + (1 - \tau_{t+1})(r_{t+1} - \delta)\right] \qquad (A6)$$

$$\frac{\mu_2}{(1 - l_{l,t})^{\gamma_2}} = \frac{\mu_1}{(C_{l,t})^{\gamma_1}} (1 - \tau_t) w_t^u \qquad (A7)$$

$$C_{l,t} + K_{l,t+1} - K_{l,t} + \phi_l K_{l,t}^2 = (1 - \tau_t)[(r_t - \delta)K_{l,t} + w_t^u l_{l,t}] + R_t \qquad (A8)$$

$$\begin{aligned} Y_{f,t} &= n^h[C_{h,t} + K_{h,t+1} - (1 - \delta)K_{h,t} + \phi_h K_{h,t}^2] + \\ &+ n^m[C_{m,t} + K_{m,t+1} - (1 - \delta)K_{m,t} + \phi_m K_{m,t}^2] + \\ &+ n^l[C_{l,t} + K_{l,t+1} - (1 - \delta)K_{l,t} + \phi_l K_{l,t}^2] + \overline{G}_t \end{aligned} \qquad (A9)$$

$$\overline{G}_t + R_t = \tau_t[Y_{f,t} - \delta(n^h K_{h,t} + n^m K_{m,t} + n^l K_{l,t})] \qquad (A10)$$

where,

$$\begin{aligned} Y_{f,t} &= A \left[\rho(A_k)^\nu (n^h K_{h,t} + n^m K_{m,t} + n^l K_{l,t})^\nu + (1 - \rho)(n^l l_{l,t})^\nu\right]^{\alpha/\nu} \times \\ &\times \left[n^h l_{h,t} + n^m l_{m,t}\right]^{1-\alpha} \end{aligned}$$

$$r_t = \frac{\alpha}{\nu} \frac{Y_{f,t} \rho \nu (A_k)^\nu (n^h K_{h,t} + n^m K_{m,t} + n^l K_{l,t})^{\nu-1}}{[\rho(A_k)^\nu (n^h K_{h,t} + n^m K_{m,t} + n^l K_{l,t})^\nu + (1 - \rho)(n^l l_{l,t})^\nu]}$$

$$w_t^s = (1 - \alpha) \frac{Y_{f,t}}{[n^h l_{h,t} + n^m l_{m,t}]}$$

$$w_t^u = \frac{\alpha}{\nu} \frac{Y_{f,t}(1 - \rho)\nu (n^l l_{l,t})^{\nu-1}}{[\rho(A_k)^\nu (n^h K_{h,t} + n^m K_{m,t} + n^l K_{l,t})^\nu + (1 - \rho)(n^l l_{l,t})^\nu]}.$$

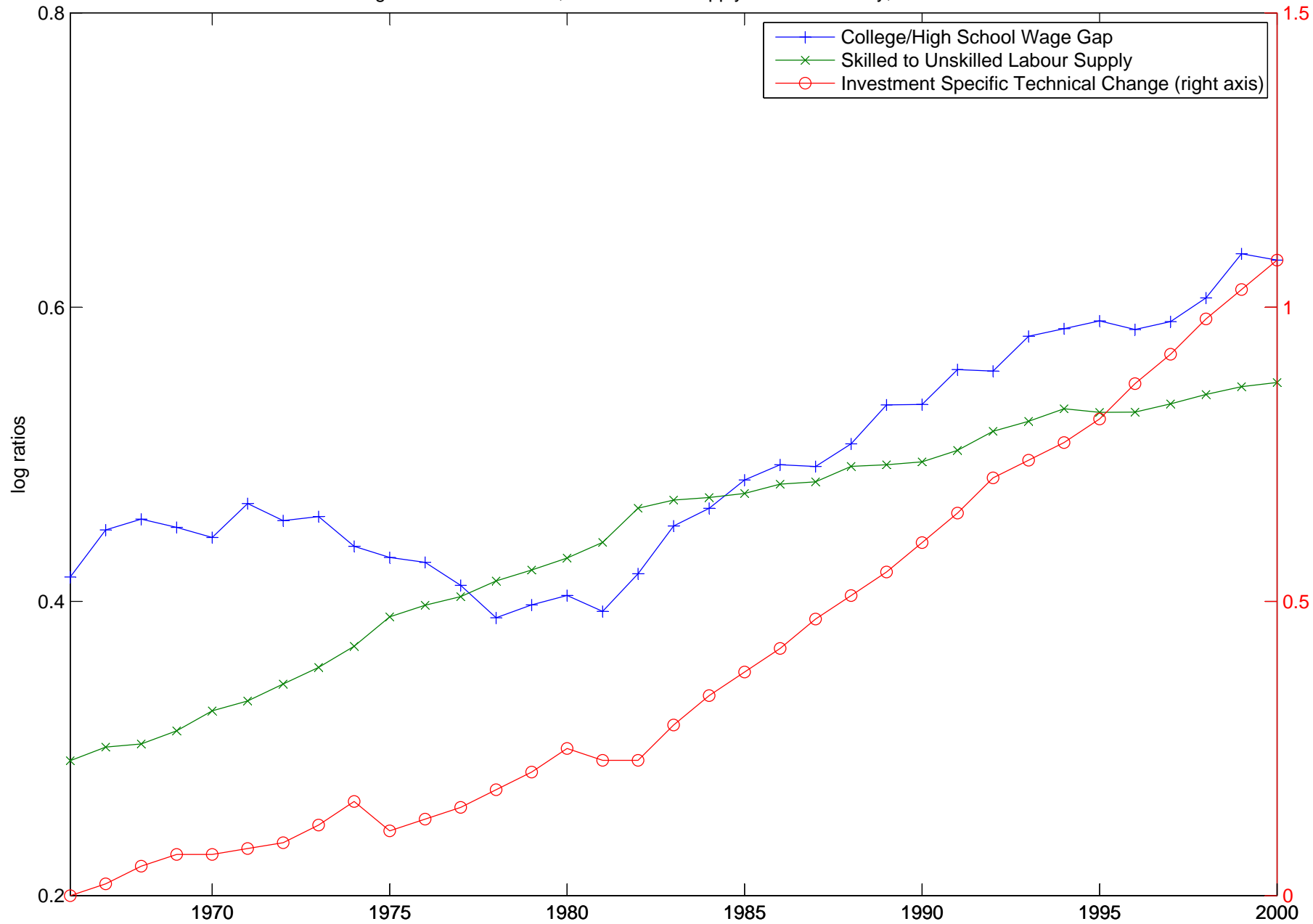Figure 1: Skill Premium, Relative Skill Supply and Productivity, 1966-2000

Figure 2: Skill Premium and the Progressivity of Tax System, 1966-2000

Legend:
- College/High School Wage Gap
- high to middle tax rate
- high to low tax rate
- middle to low tax rate

y-axis: log ratios