TASKS, TALENTS, AND TAXES*

Laurence Ales **Musab Kurnaz**
Christopher Sleet†

Tepper School of Business, Carnegie Mellon University

This Version: April 8, 2014

Abstract

This paper considers the normative implications of technical change for tax policy design. A task-to-talent assignment model of the labor market is embedded into an optimal tax problem. The impacts of technical change on wage growth across talents and the substitutability of talents across tasks emerge as key drivers of policy. The sources of technical change are measured. Evidence of polarization in the demand for tasks and a twisting of the task-talent productivity function with low talents catching up in simple tasks and falling behind in more complex ones is found. The optimal policy response is to reduce marginal income taxes at the bottom and middle of the income distribution.

*We thank seminar participants at NYU, UCSB, the IIES, the FRB Minneapolis, the Larry Jones 60th Birthday Conference and the Midwest Macro Conference (2013) for comments on earlier drafts of this paper. We thank V. V. Chari, Jonathan Heathcote, Sebastian Koehne, Hakkı Yazıcı and our discussant Marco Bassetto for very useful comments.

†Ales: ales@cmu.edu; Kurnaz: mkurnaz@andrew.cmu.edu; Sleet: csleet@andrew.cmu.edu.

1 INTRODUCTION

Technical change is inherently redistributive, complementing the labor of some while substituting for that of others. However, while the positive literature documenting the redistributive nature of technical change is extensive, normative work exploring the policy implications of such change is not.[1] Our paper fills this gap. We make theoretical and quantitative contributions. On the theoretical side, we embed a talent-to-task assignment model into an optimal tax framework. The former has been used been used by labor and trade economists to analyze the implications of technical change for the structure of wages and employment. We show how the technological parameters emphasized in this work shape optimal tax formulas. On the quantitative side, we bring a parametric assignment model to the data; we estimate the key parameters and derive the implications of technical change from the 1970's to the present day for policy. We find evidence of shifts to the task-talent productivity function: low talents catch up with high in simple tasks and fall behind in complex ones. These shifts have two effects. First, they compress wage differentials and relax incentive constraints at the bottom of the wage distribution, while expanding such differentials and tightening incentive constraints at the top. To this extent they are a force for reduced marginal taxes on low incomes and increased marginal taxes on high ones. Second, they are associated with an an increase in the comparative advantage of talented workers in complex tasks. This increase raises the sensitivity of wage differentials to tax policy and is a force for higher marginal tax rates at the bottom and lower rates at the top. Optimal policy largely depends on the balance of these forces. Under our benchmark parameterization, the first is dominant over much of the income range. The overall optimal policy response is to reduce marginal income taxes on low and middle talent workers, but leave those on higher talent workers largely intact. Transfers to low talent workers are reduced.

Positive work emphasizes the role of technological change in affecting the demand for imperfectly substitutable skills or talents. The normative literature largely abstracts from these things and, instead, focusses on the incentive to supply effort by perfectly substitutable and privately informed workers.[2] An exception is Stiglitz (1982) who allows for

[1]For a historical account of the relationship between skill and technology see Mokyr (1992), Goldin and Katz (1998), Autor, Katz, and Krueger (1998) and the references therein. Bresnahan, Brynjolfsson, and Hitt (2002) look at firm level evidence connecting technology and the demand for skills. Autor, Levy, and Murnane (2003) argue that recent technical change has led to the replacement of "routine" labor in the middle of the wage distribution. Autor, Katz, and Kearney (2006), Goos and Manning (2007) and Goos, Manning, and Salomons (2009) document "job polarization": growth in low and high skill occupations.

[2]In turn, the positive literature largely abstracts from the intensive labor supply margin.

imperfect substitutability between the effort of two different talents.[3] This assumption renders relative wages sensitive to the profile of effort across talents and, hence, tax policy. In particular, Stiglitz identifies a wage compression motive for subsidizing high and taxing low talents. By doing so the wage of high talents is compressed relative to low and the former's incentive constraints are relaxed.

We begin our analysis with a Stiglitz-type environment in which the production function is defined directly over the imperfectly substitutable labor input of many different worker types. In this setting with minimal restriction on the production function, we derive a general formula for optimal taxation. The formula provides a framework for interpreting subsequent results. Stiglitz (1982)'s wage compression channel remains operative, but now takes a more complex form: the motive to tax a given talent type k at the margin depends, in part, on the elasticity of the relative wages of all pairs of adjacent talent types (ordered by wages) with respect to k's effort. This setting suggests several ways in which technical change can influence optimal policy. First, factor augmenting technical change that is biased towards a subset of talents can do so by modifying relative wages and, hence, tightening or relaxing incentive constraints. Second, technical change that alters the effect of one talent type's effort on the relative wages of other talent types impacts policy by strengthening or diluting the wage compression channel described above. Third, Harrod neutral technical change affects policy if workers' marginal rates of substitution between consumption and earned income are altered by scalings of consumption and wages.

We next embed an assignment model into an optimal tax framework.[4] In assignment models, talented workers have a comparative advantage in complex tasks and assortative matching of workers to tasks occurs. Standard assignment models, however, omit an intensive effort margin, a societal motive for redistribution and explicitly private talent; the optimal tax framework adds these things. In the equilibria of our embedded model, workers sort themselves efficiently across tasks *conditional on the effort of other workers*. This induces an indirect production function over the effort of different talents of the sort that our earlier analysis assumed. Technological parameters that determine relative task demand and the productivity of task-talent matches in the assignment framework

[3]Related exceptions include Naito (1999), Rothschild and Scheuer (2013) and Slavík and Yazıcı (2014). The latter two are most related to our work and we discuss them below.

[4]The assignment framework is not new, originating with Roy (1950) and being extended by Sattinger (1975) and Teulings (1995). It has proven to be a rich laboratory for analyzing the role of talent-task distributions and the productivity of task-talent matches in shaping the wage distribution. Over the last few years there has been a renewed emphasis on comparative statics in this framework and on its use as a lens for viewing the implications of technical change, see Costinot and Vogel (2010), Acemoğlu and Autor (2011) and Autor and Dorn (2013).

are thus mapped to the variables and elasticities necessary for optimal tax analysis. In particular, the pattern of comparative advantage of talents across tasks shapes the sensitivity of relative wages to variations in the effort profile and, hence, policy. A local reduction in marginal taxes that induces a given talent type to increase its effort, depresses the (shadow) price of the task to which the type is assigned and, hence, its relative wage. Workers offset this by migrating into neighboring tasks, mitigating the impact on their original task's shadow price. However, the offset is partial since this migration erodes their productivity relative to neighboring talents. The greater is the comparative advantage of talented workers in complex tasks, the greater this erosion and the more sensitive are relative wages to task assignment. Thus, technical change that raises talent-complexity comparative advantage enhances the policymaker's ability to influence the wage structure through taxation. It strengthens the wage compression force identified in the more reduced form Stiglitz setting.[5]

We take our model to the data and quantify the implications of 30 years of technical change in the US for optimal policy. We treat information on occupations, incomes and hours in the Current Population Survey (CPS) as if it was generated by an equilibrium of our assignment model and use parametric assumptions and equilibrium restrictions to recover estimates of key technological parameters for the 1970's and the 2000's. To relate empirical occupations to the ordered set of tasks in our model, we order the former by the average wage paid. We recover an empirical proxy for the assignment of tasks to talents from the distribution of workers across occupations (ordered by wages). The estimation of parameters determining the demand for tasks is separated from those determining the productivity of task-talent matches by assuming a Cobb-Douglas technology for final goods as a function of tasks. This enables us to identify the demand parameters with occupational compensation shares. Parameters determining the productivity of talent-task matches and, hence, comparative advantage are recovered from the empirical assignment function and the distribution of wages across tasks using the envelope condition for wages implied by the model. After obtaining these estimates and supplementing them with calibrated preference parameters, we calculate optimal tax policies for the 1970's and 2000's.

We find evidence of relative reductions in demand for mid-level tasks and relative increases in demand for low and high level tasks. We also find evidence of a twisting

[5]Migration of workers into neighboring tasks depresses the shadow prices of these tasks inducing the talents occupying them to migrate as well. A ripple effect is created and, so, an adjustment in one talent type's effort can induce reassignment of many types, affecting their relative wages and in the process relaxing and tightening many incentive constraints. However, the greater is talent-complexity comparative advantage the more contained the impact of a policy-induced effort adjustment.

of the talent-task productivity function, with low talent productivity catching up to high talent in simple tasks and falling behind in more complex ones. The latter is associated with significant increases in the comparative advantage of more talented workers in more complex tasks. Moving from the 1970's to the 2000's, we find that under our benchmark estimation/parameterization, optimal marginal tax rates decrease on low and middle talents, but rise on higher ones.[6] Optimal transfers to workers at the first and second income deciles are reduced. The twisting of the productivity function is the main force at work. It has two effects. First, it suppresses wage variation at the bottom of the income distribution, while enhancing it at the top. This relaxes incentive constraints on low incomes, while tightening them on high ones; it is a force for reductions in optimal marginal taxes on the former and increases on the latter. These effects are slightly enhanced by the relative reduction of demand for mid-level tasks populated by mid-level talents. Second, there is a partially offsetting strengthening of the wage compression channel. Higher comparative advantage of talented workers in complex tasks increases the policymaker's motive to apply high marginal taxes on low talents. Such taxes deter low talent effort, raise low-level task prices and encourage higher talents into these tasks. The relative productivity of these task migrants is eroded, suppressing their wage premia and relaxing incentive constraints. A parallel strengthening of the policymaker's motive to reduce marginal taxes on high talents occurs. Of these two forces, the first dominates at most incomes under our benchmark parametrization.

The remainder of the paper proceeds as follows. After a brief literature review, Section 2 provides motivating facts. Section 3 gives optimal tax formulas for economies with imperfectly substitutable labor types and provides an initial discussion of the implications of technical change for policy. In Section 4 an assignment model is embedded into an optimal tax framework. An indirect production function over worker effort is derived and the parameters of the assignment model related to the relevant terms of the optimal tax formulas from Section 3. In addition, the implications of technical change for policy in a simple two talent model are discussed. Section 5 describes how the assignment model is used to identify estimates of technical change and reports these estimates. In Section 6, optimal policy for the 1970s and 2000s is computed and the implications of technical change for policy recovered. The tax formula from Section 3 is used to decompose and account for changes to optimal taxes. Section 7 concludes; appendices contain proofs and additional details.

[6] Except for those at the very top who experience reduction.

LITERATURE As noted above, our paper bridges the normative optimal taxation literature and a positive literature that analyzes the role of technical change in driving the wage distribution. Both literatures are large. Many contributions to the latter have attributed increases in the skill premium to skill-biased technical change, formalizing this insight in what Acemoğlu and Autor (2011) have called the *canonical model*, i.e. a model with two imperfectly substitutable types of workers (skilled and unskilled) and factor-augmenting technical change directed towards the skilled.[7] Recently, a more nuanced view of the labor market has emerged that recognizes that extreme (i.e. low and high wage) occupations have grown relative to middle occupations. Partly in consequence, assignment models have been increasingly adopted to analyze the joint distribution of workers across wages and occupations and the evolution of this distribution. Acemoğlu and Autor (2011) and Autor and Dorn (2013) are prominent examples.

Most contributions to the normative literature focus on labor supply. In addition to Stiglitz (1982), we mention three recent contributions that are quite related to ours and that give an (interesting) role to labor demand. Rothschild and Scheuer (2013) also develop implications for optimal tax policy in an assignment framework. In contrast to us, they focus on the two task-many talent case with, in their theory, essentially no restrictions on the distribution of talent across tasks. In this setting they provide an interesting characterization of optimal taxes and contrast such taxes with those obtained in self-confirming equilibria in which governments do not recognize their ability to influence the wage distribution. We focus on many task-many talent cases, but place much stronger restrictions on the task-talent distribution. This allows us to relate our model to those used in the positive literature to analyze technical change and to undertake comparative static exercises which Rothschild and Scheuer (2013) do not do.

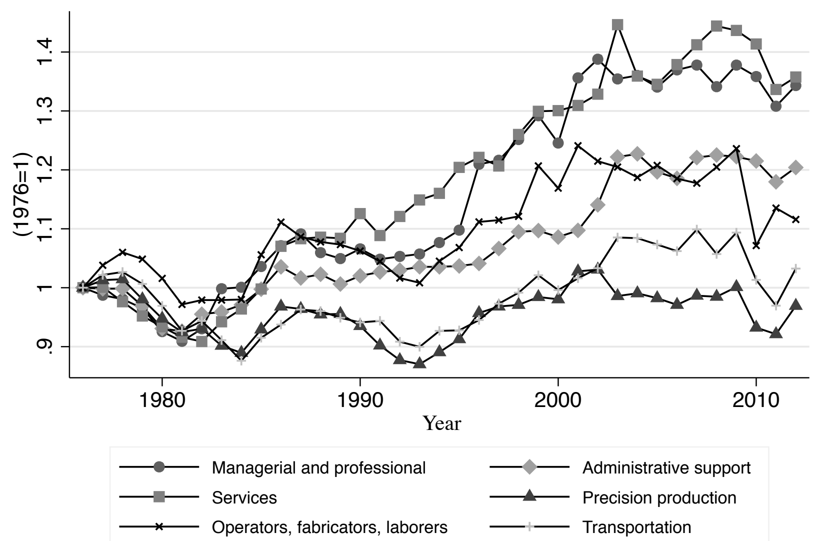
Slavík and Yazıcı (2014) apply the logic of Stiglitz (1982) to capital taxation. In their paper they introduce two sorts of capital, buildings and machines. Following the skill premium literature, they assume a machine-skill (or machine-talent) complementarity. Thus, machines raise the marginal product of the talented relative to the untalented and, as in Stiglitz (1982), this dilutes incentives. It is socially desirable to deter the accumulation of machines. In quantitative work, Slavík and Yazıcı (2014) show that this creates a rationale for quite high rates of (machine) capital taxation. Slavík and Yazıcı (2014)'s interesting contribution is complementary to ours. They endogenize technical change, which we do not, in the context of a two talent "canonical model", and develop policy implications. We treat technical change parametrically, but do so in a multi-talent/multi-task assignment setting.

[7] Examples include Acemoğlu (2002) and Krusell, Ohanian, Ríos-Rull, and Violante (2000).

Heathcote, Storesletten, and Violante (2014) analyze optimal income tax progressivity in a rich dynamic environment. They assume imperfectly substitutable skills, but do not explicitly model tasks. Our model is static, but we add assignment and, hence, endogenize the substitutability of skills and relate it to technical change. In addition, Heathcote, Storesletten, and Violante (2014) restrict optimal taxes to a parametric class, we do not.

2 Facts on Technical Change and Taxation

We first document some stylized facts that motivate our analysis. Figure 1 displays changes in average incomes across (1-digit) occupations from the 1970's to the present.[8]



Source: March CPS.

Figure 1: Evolution of average income by occupation over time.

The figure indicates considerable variation in the experience of different occupations, with some exhibiting significant average income growth and others stagnating. Moreover, occupations with slow average income growth were predominantly middle income in the 1970's, while fast growers were mainly low or high income at that time. For example, precision production, craft and repair workers had a mid-level income of $33,109 in 1975 (all incomes are expressed in 2005 dollars) and negligible income growth subsequently. In contrast, the two occupations with the fastest growing average incomes, services and managerial and professional, had average incomes in the mid-1970's of $12,912

[8]The data is taken from the March survey of the Current Population Survey (CPS). See Appendix D for additional details on the data and our sample selection.

and $40,013, placing them at opposite extremes of the income distribution. Such occupational polarization, with the middle growing more slowly than the extremes, is not confined to earnings; it is also present in various measures of occupational size and demand. Figure 2, displays changes in the share of employment of different occupations over time.[9] Here managerial/professional and service related occupations that are con-

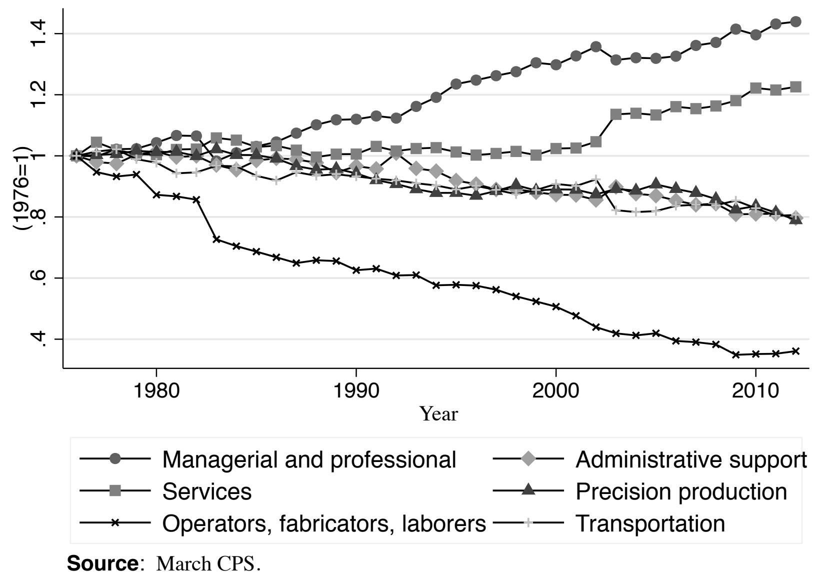


Figure 2: Evolution in size of employment by occupation over time.

centrated in the extremes of the income distribution are expanding in size, while mid-income level occupations operators and fabricators (mostly employed in manufacturing) are shrinking over time.

Overall, the picture that emerges from the CPS (and other data sources) is one in which high wage and low wage occupations are growing in size and in average compensation relative to middle ones. If talent is imperfectly substitutable across occupations, then these varied occupational fortunes suggest varied fortunes for differently-talented workers. In the remainder of the paper we consider the optimal policy response to such events. Before doing so it is interesting to document the *actual* policy response. Table 1 reports the evolution of average tax rates since the 1970's. To calculate these rates, we encode the information provided in CPS into the NBER TAXSIM simulator.[10] For each individual we use information on the year, state of residence, dependents, marital status, total and

[9] See, inter alia, Goos and Manning (2007), Goos, Manning, and Salomons (2009), Acemoğlu and Autor (2011) and Autor and Dorn (2013) for related evidence.

[10] Gouveia and Strauss (1994) emphasize the usage of effective tax rates as opposed to statutory tax rates to correctly capture the differences between market and taxable income as well as the worker's economic response to statutory tax rates.

individual income. Effective state and federal tax rates are given both with and without payroll taxes (note that we are assuming that the burden of the payroll tax falls entirely on the worker).

Table 1: Average Tax Rates on Real Labor Income.

Taxes Included	**Decade**	**Percentiles of Income**					
		10th	*25th*	*50th*	*75th*	*90th*	*99th*
State + Federal	70s	-4.9	4.6	13.1	17.9	21.5	31.5
	00s	-21.9	-6.6	7.2	14.9	20.1	34.4
State + Federal + FICA	70s	8.1	17.7	26.8	31.3	32.2	37.4
	00s	-4.8	10.6	24.9	32.1	36.8	40.7

Notes: Data on individuals is taken from the March CPS. Data for the 70s is from 1977 to 1980. Data for the 00s is from 2001 to 2010. We drop individuals with negative income and labor income below $100. Also dropped are individuals for which labor income is less than 60% of total income or more than 120% of total income. Tax rates are computed using the NBER TAXSIM calculator version 9.2. Rates reported are average tax rates on labor income of the head of household inclusive of transfer received.

Between the two time periods we observe significant reductions in average tax rates for those at low incomes, modest reductions at mid-incomes and increases at high incomes. Incorporating FICA raises average taxes by 13% to 16% at most incomes depending upon the time period, with smaller increases at the top.[11]

3 Optimal Taxation with Imperfectly Substitutable Worker Types

Mirrlees (1971)'s classic model of optimal taxation assumes that workers of different types are perfect substitutes and that final output is a weighted sum (or integral) of worker effort, with the weights given by private productivities. Stiglitz (1982) allows for a more general production function. He assumes that workers are one of two imperfectly substitutable types and interprets these types as "low" and "high" skilled. In this section, we generalize Stiglitz (1982) to K-types, but place no interpretation on a worker's type (the nature of which is defined implicitly by the production function). Our goal is to derive a general formula for optimal taxes without placing too much structure on the production

[11]The relative impact of FICA taxes on different income percentiles changes over time. This is because the maximum social security taxable earnings is approximately at the 77th percentile in the 70s, while it increases to the 90th percentile in the 00s.

function. This provides a framework for interpreting results in the assignment setting of later sections. In this general (and compared to later sections reduced form) context we briefly discuss implications of technical change for taxes.[12]

3.1 Physical Environment

Workers A continuum of workers has identical preferences over consumption $c \in \mathbb{R}_+$ and effort $e \in [0, \bar{e}]$ described by a utility function $U : \mathbb{R}_+ \times [0, \bar{e}] \to \mathbb{R}$. The function U is assumed to be concave, twice continuously differentiable on the interior of its domain, with for each $e \in [0, \bar{e}]$, $U(\cdot, e)$ increasing and for each $c \in \mathbb{R}_+$, $U(c, \cdot)$ decreasing and strictly concave. First and second partial derivatives of U are denoted U_x and U_{xy} with $x, y \in \{c, e\}$. U satisfies the Inada conditions: for all $c > 0$, $\lim_{e \downarrow 0} U_e(c, \cdot) = 0$ and $\lim_{e \uparrow \bar{e}} U_e(c, \cdot) = -\infty$. In addition, U satisfies the Spence-Mirrlees single crossing property: $-U_e(c, y/w)/\{wU_c(c, y/w)\}$ is decreasing in w.[13] Workers are partitioned across $K \in \mathbb{N}\backslash\{1\}$ "types" with a fraction π_k of workers in type group $k \in \{1, \ldots, K\}$. The fraction of workers with type less than or equal to k is denoted $\Pi_k = \sum_{j=1}^{k} \pi_j$.

Workers sell their labor to firms and pay taxes on the income that they earn. Let $T : \mathbb{R}_+ \to \mathbb{R}$ denote an income tax function. Throughout this section, income tax functions are assumed to be piecewise-differentiable with directional derivatives less than one. A worker of type k receiving wage w_k solves the problem:

$$\sup_{\mathbb{R}_+ \times [0, \bar{e}]} U(c, e) \quad \text{s.t.} \quad c \leq w_k e - T(w_k e). \tag{1}$$

Technology A representative competitive firm hires workers of all types. The firm uses a production function $F : \mathbb{R}_+^K \to \mathbb{R}_+$ defined directly on the labor inputs of the different types. The firm solves:

$$\max_{\mathbb{R}_+^K} F(e_1\pi_1, \ldots, e_K\pi_K) - \sum_{k=1}^{K} w_k \pi_k e_k,$$

where e_k is the common effort level of workers of type k. F is assumed to be a continuously differentiable, constant returns to scale function with k-th partial derivative

[12] Much of the optimal tax literature is cast in terms of a continuum of types. This literature maintains the linear production function assumption. Although versions of the results that we give below are available for continuum (of type) economies, for general constant returns to scale production functions, their derivation requires stepping outside of the framework of optimal control and maximizing an infinite-dimensional Lagrangian directly. To avoid technical complications that do not generate additional economic insight we do not do this.

[13] If, given U, consumption is a normal good, then this condition is assured.

F_k. At this stage, we place no further restrictions on F. In classical Mirrlees models $F(e_1\pi_1, \ldots, e_K\pi_K) = \sum_{k=1}^K a_k e_k \pi_k$ for some positive constants $\{a_k\}$ and workers of different types are perfectly substitutable. However, we allow for and, in this section, focus upon worker types that are imperfect substitutes in production. Since F defines what it means for a worker to be of one type or another, the economic nature of a worker's type is for the moment left implicit.

TAX EQUILIBRIUM Let G be a fixed public spending amount. Given G, a *tax equilibrium* is an income tax function $T : \mathbb{R}_+ \to \mathbb{R}$, an allocation $\{c_k, e_k\}_{k=1}^K$ and a wage profile $\{w_k\}_{k=1}^K$ such that (i) for each $k = 1, \ldots, K$, (c_k, e_k) solves (1), (ii) for each $k = 1, \ldots, K$, $w_k = F_k(e_1\pi_1, \ldots, e_K\pi_K)$ and (iii) the goods market clearing condition holds: $G + \sum_{k=1}^K c_k\pi_k \leq F(e_1\pi_1, \ldots, e_K\pi_K)$. Let $\mathscr{E}$ denote the set of tax equilibria (given G).

Remark 1. This definition restricts attention to symmetric equilibria in which all workers of a given type make the same consumption and effort choices given the tax schedule. Although non-symmetric equilibria are possible, they are not socially optimal and we exclude them. We also restrict attention to deterministic tax schedules. The literature has considered richer tax mechanisms that randomize over tax rates, see Stiglitz (1982), Brito, Hamilton, Slutsky, and Stiglitz (1995) and Hellwig (2007). The latter gives sufficient conditions for deterministic tax mechanisms to be socially optimal in utilitarian settings.

3.2 OPTIMAL POLICY

A government attaches Pareto weight g_k to workers of type k, with weights normalized to satisfy $\sum_{k=1}^K g_k = 1$. It selects a tax equilibrium to solve:

$$\sup_{\mathscr{E}} \sum_{k=1}^{K} U(c_k, e_k) g_k. \tag{PP}$$

Let T^* and $\{c_k^*, e_k^*, w_k^*\}_{k=1}^K$ denote an optimal tax equilibrium with worker types indexed so that $w_k^* = F_k(e_1^*\pi_1, \ldots, e_K^*\pi_K)$ is increasing in k.[14] . We call:

$$\tau_k^* = -\frac{U_e(c_k^*, e_k^*)}{w_k^* U_c(c_k^*, e_k^*)} - 1$$

[14]The ordering of worker types' marginal products is not given exogenously. The labeling of worker types can always be modified to reflect their marginal product ranking at the optimal allocation.

the marginal tax rate at income $q_k^* = w_k^* e_k^* > 0$.[15] The equilibrium allocation $\{c_k^*, e_k^*, w_k^*\}_{k=1}^K$ satisfies a non-binding upward incentive constraint condition if:

$$U(c_k^*, e_k^*) > U(c_{k+1}^*, q_{k+1}^*/w_k^*). \tag{NUIC}$$

Under this condition a k-type worker strictly prefers her optimal equilibrium allocation to that which she could obtain by replicating the income of a $k+1$-type worker.[16]

Our main goal is to determine how technical change, by modifying F, shapes the profile of optimal taxes. The following proposition relates optimal taxes to F and is the first step in this direction.

Proposition 1. *Let T^* and $\{c_k^*, e_k^*, w_k^*\}_{k=1}^K$ denote an optimal tax equilibrium with worker types indexed so that $w_k^* = F_k(e_1^*\pi_1, \ldots, e_K^*\pi_K)$ is increasing in k. Assume that $\{c_k^*, e_k^*, w_k^*\}_{k=1}^K$ satisfies the non-binding upward incentive condition* (NUIC)*. Optimal marginal tax rates satisfy:*

$$\frac{\tau_k^*}{1-\tau_k^*} = \frac{\Delta w_{k+1}^*}{w_{k+1}^*}\frac{1-\Pi_k}{\pi_k}\mathcal{H}_k^*\Psi_k^* + \sum_{j=1}^{K-1}\mathcal{M}_{k,j}^*\phi_{k,j}^*, \tag{2}$$

where: $\frac{\Delta w_{k+1}^*}{w_{k+1}^*} := \frac{w_{k+1}^* - w_k^*}{w_{k+1}^*}$,

$$\mathcal{H}_k^* := -\frac{\Delta_e U_c(c_k^*, e_k^*)}{U_c(c_k^*, e_k^*)}e_k^* + \frac{\Delta_e U_e(c_k^*, e_k^*)}{U_e(c_k^*, e_k^*)}\frac{w_k^*}{w_{k+1}^*}e_k^* + 1,$$

with $\Delta_e U_x(c_k^*, e_k^*) := \dfrac{U_x(c_k^*, e_k^*) - U_x\left(c_k^*, \frac{w_k^*}{w_{k+1}^*}e_k^*\right)}{e_k^* - \frac{w_k^*}{w_{k+1}^*}e_k^*}$, $x \in \{c, e\}$,

$$\Psi_k^* := \sum_{j=k}^{K-1}\mathcal{N}_{k,j}^*\left\{1 - \frac{g_{j+1}U_c(c_{j+1}^*, e_{j+1}^*)}{\pi_{j+1}\chi^*}\right\}\left(\frac{U_c(c_k^*, e_k^*)}{U_c(c_{j+1}^*, e_{j+1}^*)}\right)\frac{\pi_{j+1}}{1-\Pi_k},$$

with $\mathcal{N}_{k,j}^* := \prod_{i=k+1}^{j}\frac{U_c(c_i^*, q_i^*/w_{i+1}^*)}{U_c(c_i^*, e_i^*)}$, $\mathcal{M}_{k,j}^* := \frac{U_c(c_k^*, e_k^*)}{U_e(c_k^*, e_k^*)e_k^*}\frac{U_e\left(c_j^*, q_j^*/w_{j+1}^*\right)}{U_c(c_j^*, e_j^*)}\frac{q_j^*}{w_{j+1}^*}\Psi_j^*\frac{1-\Pi_j}{\pi_j}\frac{\pi_j}{\pi_k}$ *and* $\phi_{k,j}^* = \frac{e_k^*}{w_{j+1}^*/w_j^*}\frac{\partial w_{j+1}^*/w_j^*}{\partial e_k}(e_1^*, \ldots, e_K^*)$. *$\chi^*$ is the optimal shadow price of resources (i.e. the multiplier on the goods market clearing constraint).*

Proof. See Appendix A. □

[15] The optimal program (PP) only determines T^* at the points $\{q_k^*\}$. However, T^* may be chosen so that τ_k^* is the left derivative of T^* at q_k^*; if T^* is differentiable at q_k^*, then it is the derivative of T^*. The Inada conditions on worker utilities ensure that absent taxes workers would always choose $e \in (0, \bar{e})$. Consequently, in a tax equilibrium, a wedge between an worker's marginal rate of substitution and its wage is due to taxation and not to a boundary condition on effort.

[16] See footnote 18 for some further discussion of this condition.

Remark 2. The relative wage-effort elasticities $\phi^*_{k,j}$ play an important role in the remainder of the paper. In particular, $\phi^*_{k,k-1} = -\phi^*_{k,k}$ gives the elasticity of complementarity (and the reciprocal of the Hicks elasticity of substitution) between the k and $k-1$ types' efforts.

We now discuss and interpret the optimal tax formula in (2). This formula has two main terms, which we label "Mirrlees" and "Wage Compression":

$$\frac{\tau^*_k}{1-\tau^*_k} = \underbrace{\frac{\Delta w^*_{k+1}}{w^*_{k+1}} \frac{1-\Pi_k}{\pi_k} \mathcal{H}^*_k \Psi^*_k}_{\text{Mirrlees}} \quad + \underbrace{\sum_{j=1}^{K-1} \mathcal{M}^*_{k,j} \phi^*_{k,j}}_{\text{Wage Compression}} .$$

The "Mirrlees" term is not new: it is the discrete analogue of terms found in continuous-type models with exogenous productivity (e.g Mirrlees (1971) and Saez (2001)). However, in contrast to these models, and importantly for the subsequent analysis, the component $\frac{\Delta w^*_{k+1}}{w^*_{k+1}}$ is now endogenous. The second "Wage Compression" term only occurs in models with imperfectly substitutable labor types, such as ours and is not present in Mirrlees (1971) or Saez (2001).

Mirrlees term This term has four components.[17] $\mathcal{H}_k$ is a discrete approximation to $\frac{1+\mathscr{E}_{u,k}}{\mathscr{E}_{c,k}}$, where $\mathscr{E}_{c,k}$ and $\mathscr{E}_{u,k}$ are, respectively, the compensated and uncompensated labor supply elasticities at (c^*_k, e^*_k). If worker preferences are additively separable, this reduces to one plus (a discrete approximation to) the reciprocal of the Frisch elasticity. If preferences are further restricted to be quasi-linear in consumption, Ψ^*_k reduces to $\sum_{j=k}^{K-1} \left\{1 - \frac{g_{j+1}}{\pi_{j+1}}\right\} \frac{\pi_{j+1}}{1-\Pi_k}$, where it captures the government's redistributive motive. $\frac{1-\Pi_k}{\pi_k}$ is the reciprocal of the hazard; this plays an important role in conventional optimal tax analysis since, if types have compact support, it implies zero marginal taxes at the maximal income. In our analysis it is a parameter unaffected by technical change and, thus, plays a smaller role. In contrast, the wage growth (across types) term $\frac{\Delta w^*_{k+1}}{w^*_{k+1}}$ is important in what follows. To understand its role it is useful to recall briefly the mechanism design formulation of the government's problem. In this formulation, the government chooses a mechanism that maps reported types to incentive-compatible allocations of consumption and effort.[18] The solution to this problem gives an optimal allocation; prices and (optimal) taxes are then selected to ensure implementation of this allocation as part of a tax

[17] For detailed discussion of the role played by these components in a continuous-type setting see Salanié (2011). Saez (2001) relates the components of this term to labor supply elasticities, the type distribution and the government's redistributive motive.

[18] The formulation is fully specified in Appendix A.

equilibrium. The mechanism design problem is:

$$\sup \sum_{k=1}^{K} U(c_k, e_k) g_k$$

s.t. for each k,

$$U(c_{k+1}, e_{k+1}) \geq U\left(c_k, \frac{w_k e_k}{w_{k+1}}\right) \tag{3}$$

$$w_k = F_k(e_1\pi_1, \ldots, e_K\pi_K)$$

and

$$G + \sum_{k=1}^{K} c_k \pi_k \leq F(e_1\pi_1, \ldots, e_K\pi_K).$$

Condition (3) is the (local downward) incentive-constraint,[19] which requires that a $k+1$-th type worker is better off reporting her own type, than reproducing the income of and, hence, mimicking a k-th type worker. Crucially, the wage ratio w_{k+1}/w_k appears on the right hand side of this constraint. Higher values of this ratio reduce the effort that a $k+1$-th type worker must exert to mimic a k-type and, hence, tighten the incentive constraint and are associated with greater distortions of allocations. Thus, higher wage growth across the k and $k+1$ types is, other things equal, a force for higher marginal taxes on the k-th type.

WAGE COMPRESSION TERM The second term in (2) does *not* appear in standard optimal tax equations that are derived from models with linear production functions and exogenous wages. In settings with non-linear production functions, such as ours, the effort of the k-th worker type can affect the marginal rate of transformation and, hence, the ratio of wages between the j and $j+1$-th types. Following the logic of the previous paragraph, more compressed wage ratios relax incentive constraints and to the extent that the effort of a given type reduces or enhances such compression it should be encouraged through taxation. In particular, larger values of the relative wage-effort elasticities $\phi^*_{k,j}$ are a force

[19] Again, the ordering of types is chosen to be consistent with the ordering over optimal wages. By standard arguments, only local upwards and downwards incentive constraints potentially bind. The problem omits the local upward incentive-constraints, $U(c_k, e_k) \geq U\left(c_{k+1}, \frac{w_{k+1}e_{k+1}}{w_k}\right)$, which here and in Proposition 1 we assume to be non-binding. Later in our later numerical calculations in Section 5 we verify that such constraints do not bind. Our treatment of upward constraints follows many contributors who give tax formulas in the absence of such conditions and verify that they are non-binding in numerical calculations, see, e.g. Saez (2001) or Rothschild and Scheuer (2013). However, the proof of Proposition 1 indicates how the optimal tax formula would generalize if such constraints were to be binding.

for higher marginal taxes on the k-th type. Stiglitz (1982) identifies this wage compression channel in a two type model. In that case there is only one binding incentive constraint and $-\phi^*_{1,1} = \phi^*_{2,1} = 1/\mathcal{E}^*$, where $\mathcal{E}^*$ is the elasticity of substitution between the two worker types (i.e. $\frac{w_2/w_1}{e_2/e_1}\frac{\partial e_2/e_1}{\partial w_2/w_1}$) at the optimum. Assuming this is positive, compression of wages between the two types, requires that the effort of the high (resp. low) type should be relatively encouraged (resp. discouraged). Since the first term in (2) is zero for $k = K = 2$, this translates into an optimal marginal income subsidy for high types and an enhanced marginal income tax for low types.[20]

THE FORM OF F The functional form for F plays an important role in shaping wage growth across types $\Delta w^*_{k+1}/w^*_{k+1}$ and relative wage-effort elasticities $\phi^*_{k,j}$ and, hence, optimal taxes. In most contributions to the public finance literature F is taken to be a weighted sum of type efforts and, hence, $\Delta w^*_{k+1}/w^*_{k+1}$ and $\phi^*_{k,j}$ are treated as structural and invariant to policy. A natural and weaker alternative is to require F to be a CES function.[21] This assumption permits policy to affect $\Delta w^*_{k+1}/w^*_{k+1}$, but continues to treat the elasticity of substitution and, hence, the relative wage-effort elasticities as structural. It also places strong restrictions on the latter requiring that they equal:

$$\phi^*_{k,j} = \begin{cases} -\frac{1}{\mathcal{E}} & j = k \\ \frac{1}{\mathcal{E}} & j = k-1 \\ 0 & \text{otherwise,} \end{cases}$$

where $\mathcal{E}$ is the elasticity of substitution between the effort of worker type pairs. Thus, for each worker type k, the elasticities $\phi^*_{k,j}$ are non-zero only locally (i.e. a variation in a type's effort only affects its wage relative to others, it does not affect the relative wage of other type pairs) and all elasticities $\phi^*_{k,k}$ and $\phi^*_{k,k-1}$ take common values independent of k.[22]

TECHNICAL CHANGE The formulas in Proposition 1 point to several channels through which technical change can influence optimal policy. First, and most simply, it could raise the return to effort of all workers at a given effort profile $\{e_k\}$ through a pure Harrod

[20]The motive for this is to compress the relative wage. Moreover, this policy does not imply that high types have lower *average* taxes.

[21]See Heathcote, Storesletten, and Violante (2014) for an analysis of optimal taxation within a special class of tax functions that makes such an assumption in a rich dynamic setting.

[22] Salanié (2011) raises related concerns. He asserts: "It is, unfortunately, quite difficult to specify a production function that models the limits to factor substitution with an infinite number of factors." (Chapter 4, p.111). He emphasizes that the substitutability of similar and dissimilar worker types may be quite different, but that such differences cannot be accommodated under the CES assumption.

neutral shift. This can impact policy via the differential labor supply response of different worker types.[23] Second, the technical change could be directed towards some subset of workers. In particular, it could additively translate the function mapping the workers' effort profile $\{e_k\}$ to wage growth $\{\Delta w_{k+1}/w_{k+1}\}$ by some function. For example, if F is a CES function of the form $F(e_1\pi_1, \ldots, e_K\pi_K) = A\left[\sum_{k=1}^{K} D_k e_k^{\frac{\mathcal{E}-1}{\mathcal{E}}}\right]^{\frac{\varepsilon}{\varepsilon-1}}$, then:

$$\frac{\Delta w_{k+1}}{w_{k+1}} \approx -\log\left(\frac{w_k}{w_{k+1}}\right) = -\log\left(\frac{D_k}{D_{k+1}}\right) + \frac{1}{\mathcal{E}}\log\left(\frac{e_k}{e_{k+1}}\right)$$

and technically induced variations in the log relative CES weights $\{\log\frac{D_k}{D_{k+1}}\}$ additively translate the map from efforts to wage growth over talents. They do not affect the sensitivity of such wage growth to effort and, in particular, leave the elasticities $\phi_{k,j}$ unaltered. Such variations, by modifying the productivity of one type of worker relative to another at a given effort profile, relax or tighten incentive constraints and, hence, elicit a tax response. Third, technical change could alter the sensitivity of wages to the effort profile, i.e. it could change the functions mapping effort profiles to relative wage-effort elasticities $\phi_{k,j}$. By treating the wage distribution as exogenous and structural, the first and second channels, but not the third, can be captured. These two channels are also captured by identifying technical change with adjustments in the weights of a CES production function over type efforts. The latter treatment, as described above, is consistent with a wage compression motive via which policy affects relative wages, but it does not allow technical change to modify the strength of this motive. The assignment framework, which we adopt below, permits this.

4 Allocations with Assignment

We now consider optimal taxation in a framework with task assignment. As noted in the introduction, assignment-based frameworks have been used in the positive literature to formalize the impact of technical change on the distribution of workers across wages and occupations. As we show below they imply and, hence, micro-found an indirect production function over worker efforts. Consequently, we are able to relate key elasticities in the optimal tax equation (2) to deeper structural parameters that describe the way tasks and talent interact and the relative demand for (labor input in) tasks. We interpret changes in these parameters as technical change and conclude this section by deriving implications

[23] In our later numerical work, we close this channel down by restricting attention to utility functions of the form $\log c + h(e)$.

of such change for optimal policy in a very simple assignment model.

4.1 Physical Environment

As before workers are partitioned across types $1, \dots, K$ with a fraction π_k being of type k. We now explicitly label types as talents. In addition, there is a continuum of tasks $v \in [\underline{v}, \overline{v}]$ differentiated by complexity. A k-th talent worker faces a wage schedule $\omega_k : [\underline{v}, \overline{v}] \to \mathbb{R}_+$. The worker solves the problem:

$$\sup_{\mathbb{R}_+ \times [0,\overline{e}] \times [\underline{v},\overline{v}]} U(c, e) \quad \text{s.t.} \quad c \leq \omega_k(v)e - T(\omega_k(v)e). \tag{4}$$

The productivity of a talent k worker in task v is given by $a_k(v) \in \mathbb{R}_+$. The productivity functions $\{a_k\}$, $a_k : [\underline{v}, \overline{v}] \to \mathbb{R}_+$ satisfy the following condition.

Assumption 1. *The functions $a_k : [\underline{v}, \overline{v}] \to \mathbb{R}_+$, $k \in \{1, \dots, K\}$ are continuous and satisfy for each $k \in \{1, \dots, K-1\}$ and $v', v \in [\underline{v}, \overline{v}]$ with $v' > v$, $\log a_{k+1}(v') - \log a_k(v') \geq \log a_{k+1}(v) - \log a_k(v)$.*

The latter assumption implies that a is a weakly log super-modular function of talent and task and that higher talents have a weak comparative advantage in more complex tasks. We often strengthen this assumption by requiring a to be strictly log super-modular in task and talent: $k \in \{1, \dots, K-1\}$, and $v', v \in [\underline{v}, \overline{v}]$ with $v' > v$, $\log a_{k+1}(v') - \log a_k(v') > \log a_{k+1}(v) - \log a_k(v)$. We also sometimes supplement Assumption 1 with the absolute advantage condition $a_{k+1} > a_k$. Task output is linear in effective labor input. Let Λ_k be the distribution of talent group k across tasks with (i) $\Lambda_k([\underline{v}, \overline{v}]) = \pi_k$ and (ii) density λ_k defined at almost every v. If workers in talent group k exert effort e_k, then for almost every task v, task output is:

$$y(v) = \sum_{k=1}^{K} \lambda_k(v) a_k(v) e_k.$$

Final output Y is produced from task output $y : [\underline{v}, \overline{v}] \to \mathbb{R}_+$ using a CES-technology:

$$Y = \begin{cases} A \left\{ \int_{\underline{v}}^{\overline{v}} b(v) y(v)^{\frac{\varepsilon-1}{\varepsilon}} dv \right\}^{\frac{\varepsilon}{\varepsilon-1}} & \varepsilon \in \mathbb{R}_+ \backslash \{1\}, \\ A \exp \left\{ \int_{\underline{v}}^{\overline{v}} b(v) \ln y(v) dv \right\} & \varepsilon = 1, \end{cases} \tag{5}$$

where $A > 0$ and $b : [\underline{v}, \overline{v}] \to \mathbb{R}_{++}$ is a continuous function satisfying $B(\overline{v}) = 1$, $B(v) := \int_{\underline{v}}^{v} b(v') dv'$.

Remark 3 (Interpreting a and b). The function a captures the idea that different workers may be more or less effective at performing specific tasks or using task-specific capital. Assumption 1 implies that more complex tasks are more talent-intensive and more talented workers have a comparative advantage in these tasks. The formulation of production here follows that in the assignment literature, e.g. Costinot and Vogel (2010), with the important addition of an intensive effort margin. Note that analogous to this literature, talent refers to relative ability in complex tasks. Unless the comparative advantage assumption is supplemented with an absolute advantage condition, "talented" workers need not have greater productivity in all tasks and need not earn higher wages in equilibrium.[24]

Later we allow for the possibility that a may change over time. We interpret such change as technical progress and allow it to depend upon both worker talent and task complexity. In particular if, for each v and $k' > k$, $\log \frac{a_{k'}(v)}{a_k(v)}$ increases, then technical progress is talent-biased; if for each k and $v' > v$, $\log \frac{a_k(v')}{a_k(v)}$ increases, then it is complexity-biased and if for each $k' > k$, $v' > v$, $\log \frac{a_{k'}(v')}{a_k(v')} \Big/ \frac{a_{k'}(v)}{a_k(v)}$ increases, then it is biased towards high talent-high complexity matches. In the latter case, it enhances the comparative advantage of talent in complex tasks and reduces the substitutability of talent across tasks.

The function b weights task output in the final good aggregator. Variations in b may be interpreted as stemming from technological or preference-based variations in demand for different task outputs. We do not explicitly model capital. However, the model may be extended in this direction, in which case the production functions in (5), under the assumption $B(\overline{v}) \in (0, 1)$, can be reinterpreted as indirect production functions for labor across tasks after the substitution of optimal capital. The parameter $b(v)$ is then interpreted as the sensitivity of final output with respect to the labor input in task v. It is influenced not only by variations in demand for different tasks, but also variations in the capital/labor intensity of tasks. Such variations are stressed by Acemoğlu and Autor (2011) who emphasize the automatization of middle complexity tasks. A further possibility is that b captures the extent to which workers purchase task output in domestic markets, produce it at home or purchase it in foreign markets. Shifts in b for some tasks may reflect the substitution of market for home production as in Buera and Kaboski (2012) or domestic for foreign production as in Grossman and Rossi-Hansberg (2008).

A representative firm hires workers of all talents to perform tasks and combines task

[24] The assignment literature refers to a worker's innate productive attribute as "skill". Since skills are endogenous, we prefer the word talent. Our model could be reinterpreted as one in which workers exert effort partly or wholly in acquiring skills rather than working.

output to produce final output.[25] The firm pays a wage $\omega_k(v)$ per unit of effort to a worker of talent k in task v and solves:

$$\max_{\lambda,e} A\left\{\int_{\underline{v}}^{\overline{v}} b(v)\left\{\sum_{k=1}^{K}\lambda_k(v)a_k(v)e_k\right\}^{\frac{\varepsilon-1}{\varepsilon}}dv\right\}^{\frac{\varepsilon}{\varepsilon-1}} - \int_{\underline{v}}^{\overline{v}}\sum_{k=1}^{K}\omega_k(v)\lambda_k(v)e_k dv. \tag{6}$$

4.2 Tax Equilibria and the Government's Policy Problem

In the assignment setting, the definition of a tax equilibrium is modified as follows.[26]

Tax Equilibrium Let G be a fixed public spending amount. Given G, a *tax equilibrium* is an income tax function $T : \mathbb{R}_+ \to \mathbb{R}$, an allocation $\{c_k, e_k, \lambda_k\}_{k=1}^K$ and a wage profile $\{\omega_k\}_{k=1}^K$ such that (i) for each $k = 1, \ldots, K$, (c_k, e_k) and v in the support of Λ_k solves the k-th worker's problem at T and ω_k, (ii) $\{\lambda_k, e_k\}$ solves (6) at $\{\omega_k\}$, (iii) the final goods market clears:

$$G + \sum_{k=1}^{K} c_k\pi_k \leq A\left\{\int_{\underline{v}}^{\overline{v}} b(v)\left\{\sum_{k=1}^{K}\lambda_k(v)a_k(v)e_k\right\}^{\frac{\varepsilon-1}{\varepsilon}}dv\right\}^{\frac{\varepsilon}{\varepsilon-1}}, \tag{7}$$

and (iv) the labor markets clear, for all $k = 1, \ldots, K$,

$$\pi_k = \int_{\underline{v}}^{\overline{v}} \lambda_k(v)dv. \tag{8}$$

Again, let $\mathscr{E}$ denote the set of tax equilibria. Proposition 2 below characterizes tax equilibria. It contains the simple, but important implication that *conditional on effort* assignment in a tax equilibria is efficient.

[25] The organization of the production process could be further disaggregated into final goods producers who combine task output into final output and intermediate producers who use labor to produce a single task and sell their output on an intermediate tasks market. Alternatively, producers of tasks could sell directly to consumers with aggregation of task output an aspect of worker preferences. These alternatives are irrelevant for the results and questions we focus upon.

[26] As before, we constrain the set of mechanisms available to the government to ones that deterministically condition upon worker incomes. This assumption is standard in the literature and to a first approximation describes current tax codes. In our setting, it implies that the government cannot observe the task a worker performs or the amount of task output. The former may reasonably reflect the inherent difficulties in distinguishing between a worker's formal job description and the tasks that the worker actually performs. The latter assumption is relaxed in Ales, Kurnaz, and Sleet (2013), where taxation of task output is permitted. This introduces a motive for indirect taxation similar to Naito (1999). We also limit attention to tax equilibria in which workers of a given type work the same amount and workers of a given type distribute themselves across tasks according to a density.

Proposition 2. *Let* $\{c_k, e_k, \lambda_k\}_{k=1}^K$ *and* $\{\omega_k\}_{k=1}^K$ *be, respectively, the allocation and wage profile of a tax equilibrium. Then there is a tuple of threshold tasks* $\{\tilde{v}_k\}_{k=1}^{K-1}$ *such that:*

$$\lambda_k(v) = \begin{cases} 0 & v \in [\underline{v}, \tilde{v}_{k-1}) \cup (\tilde{v}_k, \overline{v}] \\ \frac{b(v)^\varepsilon a_k(v)^{\varepsilon-1}}{B_k(\tilde{v}_{k-1}, \tilde{v}_k)^\varepsilon} \pi_k & v \in (\tilde{v}_{k-1}, \tilde{v}_k), \end{cases}$$

where $B_k(\tilde{v}_{k-1}, \tilde{v}_k) := \left[\int_{\tilde{v}_{k-1}}^{\tilde{v}_k} b(v)^\varepsilon a_k(v)^{\varepsilon-1} dv\right]^{\frac{1}{\varepsilon}}$*. All workers of talent k earn a common wage* $w_k = \omega_k(v)$*,* $v \in [\tilde{v}_{k-1}, \tilde{v}_k]$*. Relative wages are given by:*

$$\frac{w_{k+1}}{w_k} = \frac{a_{k+1}(\tilde{v}_k)}{a_k(\tilde{v}_k)} = \frac{B_{k+1}(\tilde{v}_k, \tilde{v}_{k+1}) \Big/ \{\pi_{k+1} e_{k+1}\}^{\frac{1}{\varepsilon}}}{B_k(\tilde{v}_{k-1}, \tilde{v}_k) \Big/ \{\pi_k e_k\}^{\frac{1}{\varepsilon}}}. \tag{9}$$

Conditional on the effort profile $\{e_k\}$*, the equilibrium allocation of talent to tasks maximizes output and is efficient.*

Proof. See Appendix B. □

Efficiency of assignment *conditional on effort* implies that output is given by the following indirect production function over efforts:

$$F(\pi_1 e_1, \ldots, \pi_K e_K) = \sup \left\{ A \left\{ \sum_{k=1}^K B_k(\tilde{v}_{k-1}, \tilde{v}_k) \{e_k \pi_k\}^{\frac{\varepsilon-1}{\varepsilon}} \right\}^{\frac{\varepsilon}{\varepsilon-1}} \middle| \text{s.t. } \underline{v} \le \tilde{v}_1 \ldots \le \tilde{v}_{K-1} \le \overline{v} \right\}. \tag{10}$$

With F determined in this way, the environment effectively reduces to that in Section 3 and the government's problem to (PP). Now, however, the function F is micro-founded; changes in parameters of this production function can be related to changes in the demand for tasks b and the productivity of task-talent matches $\{a_k\}$. Evaluation of F at a given effort profile $\{e_k\}$ requires the solution of the problem in (10). This is an assignment problem identical to those considered in Teulings (1995), Costinot and Vogel (2010) and Acemoğlu and Autor (2011) (with the important distinction that the labor allocation is selected as part of an optimal tax equilibrium rather than being pinned down parametrically).[27] Solving the assignment problem at an effort profile $\{e_k\}$ reduces to finding a

[27] In fact the analysis on p. 758-60 of Costinot and Vogel (2010) in which the labor input across "skills" is changed in particular ways represents a partial exploration of the indirect production function.

sequence of task thresholds $\{\tilde{v}_k\}_{k=1}^{K-1}$ satisfying the discrete boundary value problem:

$$\frac{a_{k+1}(\tilde{v}_k)}{a_k(\tilde{v}_k)} = \frac{B_{k+1}(\tilde{v}_k, \tilde{v}_{k+1}) \Big/ \{\pi_{k+1} e_{k+1}\}^{\frac{1}{\varepsilon}}}{B_k(\tilde{v}_{k-1}, \tilde{v}_k) \Big/ \{\pi_k e_k\}^{\frac{1}{\varepsilon}}}, \tag{11}$$

with $\tilde{v}_0 = \underline{v}$ and $\tilde{v}_K = \overline{v}$.

Proposition 1 identifies relative wage-effort elasticities $\phi_{k,j}$ as key determinants of the wage compression channel and, hence, marginal taxes. Before turning to their determination in the assignment setting, we first deal with a difficulty in applying Proposition 1. This proposition requires workers to be ordered by their (optimal) wage. In the current section, however, they are indexed and ordered by their talent, i.e. by their ability to do complex tasks. Proposition 2 implies that the two orderings are identical in a tax equilibrium with task thresholds $\{\tilde{v}_k^*\}_{k=1}^{K-1}$ if and only if for $k = 1, \ldots, K-1$, $\frac{a_{k+1}(\tilde{v}_k^*)}{a_k(\tilde{v}_k^*)} > 1$. The latter condition is satisfied if Assumption 1 is supplemented with the absolute advantage condition, for all $k = 1, \ldots, K-1$ and $v \in [\underline{v}, \overline{v}]$, $a_{k+1}(v) > a_k(v)$. Of course, absolute advantage is stronger than is needed. In the remainder of this paper, we will *only* consider equilibria in which the two orderings are identical and, in particular, we will assume they are identical in all optimal equilibria.

If each $\log(a_{j+1}/a_j)$ is differentiable, then in a tax equilibrium, with wages increasing in talent, the elasticity terms $\phi_{k,j}$ from the optimal tax formulas in Proposition 1 can be expressed as:

$$\phi_{k,j} = -\frac{\partial \log(w_{j+1}/w_j)}{\partial \log e_k} = \begin{cases} -\frac{\partial \log(a_{j+1}/a_j)}{\partial \log \tilde{v}_j} \prod_{l=k}^{j-1} \left(\frac{\partial \log \tilde{v}_{l+1}}{\partial \log \tilde{v}_l}\right) \frac{\partial \log \tilde{v}_k}{\partial \log e_k} & j \geq k \\ -\frac{\partial \log(a_{j+1}/a_j)}{\partial \log \tilde{v}_j} \prod_{l=j}^{k-2} \left(\frac{\partial \log \tilde{v}_l}{\partial \log \tilde{v}_{l+1}}\right) \frac{\partial \log \tilde{v}_{k-1}}{\partial \log e_k} & j < k \end{cases}. \tag{12}$$

Thus, $\phi_{k,j}$ depends upon the local comparative advantage of talents j and $j+1$, the sensitivity of the $k-1$ or k-th task threshold to the effort of the k-th talent and the sensitivity of thresholds intermediate between j and k to one another. Only under very special conditions is the induced production function F a CES function. Once such case occurs when for each k and v, $a_k(v) = \alpha_k$, for a sequence of positive (and increasing) constants α_k. Then, each $\frac{\partial \log(a_{j+1}/a_j)}{\partial \log \tilde{v}_j} = 0$ and $\phi_{k,j} = 0$, talents are perfectly substitutable across tasks and F is linear. Another, although not one consistent with talent-complexity comparative advantage except when $K = 2$,[28] occurs when the a_k functions are indicators for the subintervals $[\underline{v}, \tilde{v}_1]$, $(\tilde{v}_1, \tilde{v}_2], \ldots,$ $(\tilde{v}_{K-1}, \overline{v}]$. Then workers are as substitutable as the tasks into which they are locked.

[28]Nor with smoothness or continuity of the a_k functions.

For more general cases, however, relative wage-effort elasticities are complicated functions of technological parameters and the effort profile $\{e_k\}$ and, hence, indirectly policy. Thus, they are not structural. In the appendix, we prove:

Lemma 1. *Each* $\frac{\partial \log \tilde{v}_j}{\partial \log \tilde{v}_{j+1}}$, $\frac{\partial \log \tilde{v}_{j+1}}{\partial \log \tilde{v}_j}$ *and* $\frac{\partial \log \tilde{v}_k}{\partial \log e_k}$ *is positive; each* $\frac{\partial \log \tilde{v}_{k-1}}{\partial \log e_k}$ *is negative. If* $\frac{\partial \log(a_{j+1}/a_j)}{\partial \log \tilde{v}_j} > 0$, *then* $\phi_{k,j} < 0$ *if* $j \geq k$ *and* $\phi_{k,j} > 0$ *if* $j < k$. *In addition,* $\phi_{k,k} \in [-1/\varepsilon, 0]$ *and* $\phi_{k,k-1} = [0, 1/\varepsilon]$.

Proof. See Appendix B. □

The economics behind Lemma 1 is straightforward. Consider a small increase in e_k (perhaps in response to a policy change). This raises output in tasks $[\tilde{v}_{k-1}, \tilde{v}_k]$, placing downward pressure on $[\tilde{v}_{k-1}, \tilde{v}_k]$-shadow prices and, hence, the wages of talent k workers. These workers respond by populating tasks that are both below $\tilde{v}_{k-1}$ and above $\tilde{v}_k$, thus moderating the impact of the increase in e_k on their wages. However, the impact is not fully offset: as $\tilde{v}_{k-1}$ falls and $\tilde{v}_k$ rises, k-talents move into tasks in which they have a comparative disadvantage relative to, respectively, $k-1$ and $k+1$-talent workers. Thus, w_k/w_{k-1} falls and w_{k+1}/w_k rises. Moreover, as k-talents spill into neighboring tasks, output of these tasks increases, depressing their shadow prices and inducing neighboring talents to migrate into new tasks. Workers of talent $k+1$ move into tasks above $\tilde{v}_{k+1}$, while workers of talent $k-1$ talents move into tasks below $\tilde{v}_{k-1}$. A ripple effect is created with each task threshold $\tilde{v}_j$ above k rising and each threshold below k falling. Since relative wages between adjacent talents are determined by productivity ratios at thresholds (i.e. by $a_{j+1}(\tilde{v}_j)/a_j(\tilde{v}_j)$), an effort change by talent k workers can affect relative wages across the whole spectrum of talents and be a motive for encouraging or discouraging that effort's talent.

Expressions for the threshold elasticities $\frac{\partial \log \tilde{v}_j}{\partial \log \tilde{v}_{j+1}}, \frac{\partial \log \tilde{v}_{j+1}}{\partial \log \tilde{v}_j}$ and $\frac{\partial \log \tilde{v}_k}{\partial \log e_k}$ are given in the proof of Lemma 1. They point to the role of the parameters b and a in influencing the sensitivity of task assignment and, hence, relative wages to a given talent's effort. Suppose that workers of talent $j-1$ encroach on the tasks performed by talent j workers. The latter will respond by migrating into more complex tasks. If there is much demand and, hence, high b-values for tasks immediately above $\tilde{v}_j$, then these tasks will absorb this migration and $\tilde{v}_j$ will change little. Conversely, if b-values in this neighborhood are low, then talent j-workers will migrate further up through the task set. In the former case, the impact on the w_{j+1}/w_j wage differential will be muted; in the latter case, it will be enhanced. The a function and task-talent complementarity plays a dual role. Greater comparative advantage of high talents in complex tasks reduces the sensitivity of thresholds to a given talent's effort: workers are less substitutable across tasks. On the other hand, relative wages are much more sensitive to any threshold adjustment that does occur.

4.3 An Example: Technical Change in the Two Talent Model

In this section we use a simple two talent-many task assignment framework to highlight implications of technical change for tax policy.

Physical Environment The number of talents is restricted to $K = 2$ which, for concreteness, we label low (L) and high (H). Low and high talents have productivities $a_L(v)$ and $a_H(v)$ in task v, with, as before, the function a weakly log super-modular in $k \in \{L, H\}$ and $v \in [\underline{v}, \overline{v}]$. Preferences are restricted to be quasi-linear in consumption:

$$U(c, e) = c - \frac{e^{1+\gamma}}{1+\gamma},$$

where $\gamma > 0$. The government's objective is Paretian with weights g_k, $k \in \{L, H\}$ and to ensure that the government is motivated to redistribute from high to low talents: $g_L > \pi_L$.

In this setting, the Mirrlees and Wage Compression components of tax formulas can be combined. Substituting for the "talent premium" $\frac{w_H^*}{w_L^*} = \frac{a_H(\tilde{v}^*)}{a_L(\tilde{v}^*)}$ and relative incomes $\frac{q_H^*}{q_L^*} = \frac{a_H(\tilde{v}^*)}{a_L(\tilde{v}^*)}\frac{e_H^*}{e_L^*}$, the optimal marginal tax equations (2) reduce to:

$$\frac{\tau_L^*}{1-\tau_L^*} = \left(\frac{g_L}{\pi_L} - 1\right)\left\{1 - \left(\frac{a_L(\tilde{v}^*)}{a_H(\tilde{v}^*)}\right)^{1+\gamma}\left\{1 - \frac{1}{\mathcal{E}^*}\right\}\right\} \qquad \geq 0 \qquad (13)$$

and

$$\frac{\tau_H^*}{1-\tau_H^*} = \left(\frac{g_H}{\pi_H} - 1\right)\left(\frac{a_L(\tilde{v}^*)e_L^*}{a_H(\tilde{v}^*)e_H^*}\right)^{1+\gamma}\frac{1}{\mathcal{E}^*} \qquad \leq 0, \qquad (14)$$

where $\mathcal{E}^* = \mathcal{E}(\tilde{v}^*; a, b)$ denotes the value of the elasticity of substitution at the optimum and this is given by:

$$\mathcal{E}(\tilde{v}; a, b) := -\frac{\partial \log \frac{e_H}{e_L}}{\partial \log \frac{w_H}{w_L}} = \varepsilon + \frac{1}{\frac{\partial \log a_H/a_L}{\partial v}(\tilde{v})}\left[\frac{b_H(\tilde{v})}{B_H(\tilde{v})} + \frac{b_L(\tilde{v})}{B_L(\tilde{v})}\right] \geq \varepsilon, \qquad (15)$$

with $B_L(\tilde{v}) := \int_{\underline{v}}^{\tilde{v}} b(v)^\varepsilon a_L(v)^{\varepsilon-1} dv$, $B_H(\tilde{v}) := \int_{\tilde{v}}^{\overline{v}} b(v)^\varepsilon a_H(v)^{\varepsilon-1} dv$ and for $k \in \{L, H\}$, $b_k(\tilde{v}) := b(\tilde{v})^\varepsilon a_k(\tilde{v})^{\varepsilon-1}$. Equations (13) to (15) indicate the role of $a = (a_L, a_H)$ and b in shaping taxes. Technical change that induces increases in the optimal relative wage $\frac{a_H(\tilde{v}^*)}{a_L(\tilde{v}^*)}$ either directly or indirectly through an increase in the threshold task is associated with a rise in marginal taxes on low talents and a fall in marginal subsidies on high talents. Similarly, technical change that is associated with a reduction in the optimal value of the elasticity of substitution $\mathcal{E}^*$ is associated with a rise in marginal taxes on low talents and a rise in marginal subsidies on high talents. Of course, while changes in a and b directly affect the functions $\frac{a_H}{a_L}(\cdot)$ and $\mathcal{E}(\cdot; a, b)$, they also prompt adjustments in relative efforts and the threshold task (as does the policy response itself). These sometimes offsetting adjustments complicate analysis of the impact of technical change on $\frac{a_H(\tilde{v}^*)}{a_L(\tilde{v}^*)}$ and $\mathcal{E}^*$ and, hence, taxes.

To proceed further, combine the workers' equilibrium first order conditions:

$$\left(\frac{1-\tau_H}{1-\tau_L}\right)\frac{a_H(\tilde{v})}{a_L(\tilde{v})} = \left(\frac{e_H}{e_L}\right)^{\gamma}$$

with the relative wage condition (9) to get:

$$\frac{w_H}{w_L} = \frac{a_H(\tilde{v})}{a_L(\tilde{v})} = \left(\frac{B_H(\tilde{v})}{B_L(\tilde{v})}\right)^{\frac{\gamma\varepsilon}{1+\gamma\varepsilon}} \left(\frac{\pi_L}{\pi_H}\right)^{\frac{\gamma}{1+\gamma\varepsilon}} \left(\frac{1-\tau_H}{1-\tau_L}\right)^{\frac{1}{1+\gamma\varepsilon}}.$$

The latter shows explicitly how the a and b functions and relative marginal taxes determine the threshold and, hence, relative wages. Substituting for optimal marginal taxes from (13) and (14) gives:

$$\frac{a_H(\tilde{v}^*)}{a_L(\tilde{v}^*)} = \left(\frac{B_H(\tilde{v}^*)}{B_L(\tilde{v}^*)}\right)^{\frac{\gamma\varepsilon}{1+\gamma\varepsilon}} \left(\frac{\pi_L}{\pi_H}\right)^{\frac{1+\gamma}{1+\gamma\varepsilon}} \times \left\{ \frac{g_H - (\pi_H - g_H)\left(\frac{a_H(\tilde{v}^*)}{a_L(\tilde{v}^*)}\right)^{(1+\gamma)(\varepsilon-1)} \left(\frac{B_H(\tilde{v}^*)}{B_L(\tilde{v}^*)}\right)^{-(1+\gamma)\varepsilon} \frac{1}{\mathcal{E}^*}}{g_L - (g_L - \pi_L)\left(\frac{a_L(\tilde{v}^*)}{a_H(\tilde{v}^*)}\right)^{1+\gamma}\left(\frac{\mathcal{E}^*-1}{\mathcal{E}^*}\right)} \right\}^{\frac{1}{1+\gamma\varepsilon}}. \quad (16)$$

It follows easily from (16) that if $\varepsilon \geq 1$ (so that goods, and, hence, efforts of different talents are gross substitutes) and if $\mathcal{E}(\cdot; a, \cdot)$ is (locally) constant with respect to small variations in $\tilde{v}$ and small complexity-biased perturbations of b that raise B_H/B_L, then the latter lead to increases in $\tilde{v}^*$ and w_H^*/w_L^*. Increases in the relative demand for more complex tasks raise the relative shadow price of such tasks and encourage less talented workers to migrate into them ($\tilde{v}^*$ rises). However, such task-upgrading erodes the comparative advantage of low talents; the talent premium rises. These effects are mitigated by adjustments in relative efforts that occur in response to wage adjustments and that are reinforced by changes to tax policy. In particular, from (13), marginal taxes on low talents rise suppressing effort of these types and, hence, the rise in the talent premium.

The preceding discussion supposed that the elasticity of substitution $\mathcal{E}^*$ is constant in response to a shifts in task demand; in general it is not. It may rise or fall as a direct effect of the change in b or the indirect effect of changes in $\tilde{v}^*$ on $\frac{1}{\frac{\partial \log a_H/a_L}{\partial v}(\tilde{v}^*)}\left\{\frac{b_H(\tilde{v}^*)}{B_H(\tilde{v}^*)}+\frac{b_L(\tilde{v}^*)}{B_L(\tilde{v}^*)}\right\}$. These changes modify the sensitivity of the log productivity ratio $\log(a_H/a_L)$ and the log weight ratio $\log(B_H/B_L)$ to $\tilde{v}^*$ and, hence, the sensitivity of the talent premium to relative efforts. They may act to reinforce or offset the responses just described. To the extent that $\mathcal{E}^*$ is increased, the government is encouraged to reduce relative taxation of low talents and to permit a further increase in the talent premium. The reverse is true if $\mathcal{E}^*$ falls.

Now, assume that $a_H(v) = a_L(v)\exp\{\alpha_1 + \alpha_2(v - \underline{v})\}$ so that α_1 controls the absolute advantage of high talents (in the lowest task) and α_2 controls their comparative advantage in more complex tasks. If, as before $\varepsilon > 1$ and $\frac{b_H(\tilde{v}^*)}{B_H(\tilde{v}^*)}+\frac{b_L(\tilde{v}^*)}{B_L(\tilde{v}^*)}$ is locally constant, then small technologically induced increases in α_2 will, from (16), both raise the talent premium and reduce the elasticity of substitution $\mathcal{E}^*$.[29] Low talent marginal taxes τ_L^* will rise both because $\frac{a_H}{a_L}(\tilde{v}^*)$ rises and because the wage compression channel is enhanced via the reduction in $\mathcal{E}^*$: as workers become less substitutable, the government is encouraged to offset the rise in the talent premium by discouraging low talent effort through taxation. Increases in α_1 work in a related way, but absent any reinforcing adjustment in $\mathcal{E}^*$. As in the case of complexity-biased perturbations in the b functions, adjustments in the $\frac{b_H(\tilde{v}^*)}{B_H(\tilde{v}^*)}+\frac{b_L(\tilde{v}^*)}{B_L(\tilde{v}^*)}$ term (either direct through changes to the b_k functions or indirect through adjustments to $\tilde{v}^*$) may work to reinforce or dampen these effects.

SUMMARY Technical change that increases the talent wage premium and reduces the substitutability of talents is associated with higher optimal marginal taxes on low talents. Change that increases the talent income premium and the substitutability of talents is associated with lower marginal subsidies on high talents. In general, the technical parameters a and b influence both talent premia and talent substitutability with the latter a fairly complicated function of parameters and endogenous task assignment. The analysis is still more complicated in settings with multiple talents. Such settings seem essential, however, for exploring the implications of recently documented changes in the patterns of wage premia across talents and skills. In particular, to capture the policy implications of technologically-driven job polarization, a model with more than two talents is required.

[29] In this case the task threshold $\tilde{v}^*$ falls: the increased productivity of high talents in complex tasks reduces the relative shadow price of such tasks and encourages high talents to downgrade their tasks. Despite some erosion of their comparative advantage, their relative wages rise.

5 Measuring Technical Change

In this section, we measure the extent of technical change in the US. Our main data source is the Current Population Survey (CPS).[30] We proceed as if this data were generated by a (sub-optimal) tax equilibrium and use parametric assumptions and equilibrium restrictions from our model to identify and estimate the technological parameters a and b in the 1970's and 2000's. In Section 6, we calculate optimal tax equilibria at these estimated parameters.

5.1 Determining Types and Tasks

Mapping empirical occupations to ordered sets of tasks The CPS categorizes workers into $M = 302$ distinct occupations; it also provides information on worker earnings and hours worked from which a measure of wages can be imputed. Our model involves an interval of tasks ordered by complexity. To relate empirical occupations to modeled tasks, we utilize the model's implication that equilibrium wages are rising in task complexity. We normalize the task space to $[\underline{v}, \overline{v}] = [0, 1]$ and sub-divide this interval into M sub-intervals of length $\Delta v = \frac{1}{M}$, $\mathcal{V}_m = [v_{m-1}, v_m]$. We calculate the imputed average wage in each occupation using 1970's data and rank occupations according to this wage. The m-th ranked occupation is then mapped to the m-th subinterval $\mathcal{V}_m$.[31]

Recovering the empirical assignment function $\tilde{v}$ The model in Section 4 featured a finite number of talents; this facilitated the derivation of analytical results. However, for the remainder of the paper we find it convenient to treat worker talent symmetrically with task complexity and to assume that workers are distributed uniformly across an interval of talents, $k \in [\underline{k}, \overline{k}]$.[32] Thus, a worker's talent should now be an interpreted as an index (and a rank), the implications of which for productivity are captured by a function: $a : [\underline{k}, \overline{k}] \times [\underline{v}, \overline{v}] \to \mathbb{R}_+$. In particular, we shall restrict a so that higher talent-indexed workers have a comparative advantage in higher complexity-indexed tasks (occupations).[33] The set $[\underline{k}, \overline{k}]$ is normalized to $[0, 1]$.

The continuous analogue of the task thresholds $\{\tilde{v}_k\}$ is a task assignment function:

[30] Further details of our use and treatment of the data are given in Appendix D.

[31] This approach is essentially that taken by Acemoğlu and Autor (2011).

[32] The convenience is two fold. First, since occupational (task) data is discrete, assuming a continuous set of talents avoids having to deal with talent groups that are distributed across adjacent occupations. Second, it allows us to apply numerical optimal control methods to solve the problem.

[33] Note that although the distribution over the (ordinal) talent index is uniform, the distribution over (cardinal) productivities is not. It is induced endogenously by a and by the assignment of talent to tasks.

$\tilde{v} : [\underline{k}, \overline{k}] \to [\underline{v}, \overline{v}]$. This function is strictly increasing in our model; denote its inverse by $\tilde{k}$. Under the assumption that workers are distributed uniformly across talent indices, $\tilde{k}$ is the distribution of workers across tasks. Consequently, we treat the distribution of workers across ordered occupations as the empirical counterpart of $\tilde{k}$ and $\tilde{v}$ to be the inverse of this. A formal statement of the continuous talent-continuous task model can be found in Appendix C.

5.2 ESTIMATING b

To facilitate identification of the a and b functions, we restrict the elasticity of substitution between task outputs to be one ($\varepsilon = 1$). Then $b(v)$ is simply the elasticity of final output with respect to task v output and, hence, the share of total compensation paid to workers in task v. Thus, estimates of b may be be calculated from compensation shares independently of knowledge of the a's. Specifically, under the Cobb-Douglas restriction, the firm's first order conditions from the continuous-talent version of (6), imply for almost all (k, v):

$$\omega(k, v) = Y \frac{a(k, v) b(v)}{y(v)}. \tag{17}$$

In the continuous talent setting, task output is given by $y(v) = a(\tilde{k}(v), v) e(\tilde{k}(v)) \tilde{k}_v(v)$, with $\tilde{k}_v$ the derivative of $\tilde{k}$. Combining this with (17) and integrating over $\mathcal{V}_m$ gives total labor income in occupation m in terms of the b-function:

$$\int_{\mathcal{V}_m} \omega(\tilde{k}(v), v) e(\tilde{k}(v)) \tilde{k}_v(v) dv = Y \int_{\mathcal{V}_m} b(v) dv.$$

Average income in occupation m, i_m, is then obtained by dividing both sides by the mass of workers in the occupation, S_m:

$$i_m := \frac{1}{S_m} \int_{\mathcal{V}_m} \omega(\tilde{k}(v), v) e(\tilde{k}(v), v) \tilde{k}_v(v) dv = \frac{Y}{S_m} \int_{\mathcal{V}_m} b(v) dv.$$

Thus, the average value of b in occupation m, $b_m \Delta v := \int_{v_{m-1}}^{v_m} b(v) dv$, is:

$$b_m = \frac{S_m i_m}{\Delta v Y}, \qquad \forall\ m = 1, \ldots, M. \tag{18}$$

We identify Y with per capita labor income.[34] A smooth estimate of the b-function is obtained by fitting a LOWESS model to $\{v_m, \log b_m\}$ data.[35] Figure 3 displays estimates of b for the 1970's and the 2000's. The figure shows that b rises (slightly) for low and (sig-

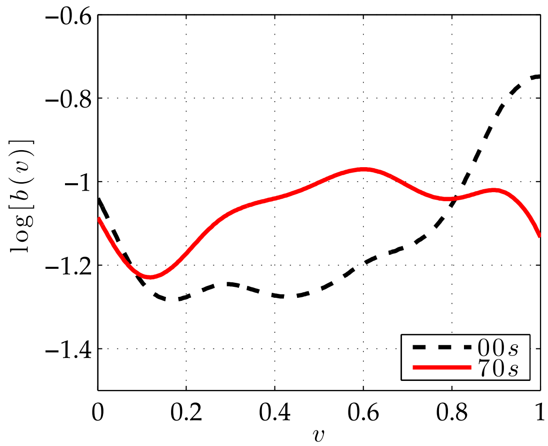


Figure 3: Evolution of $\log(b(v))$ across decades.

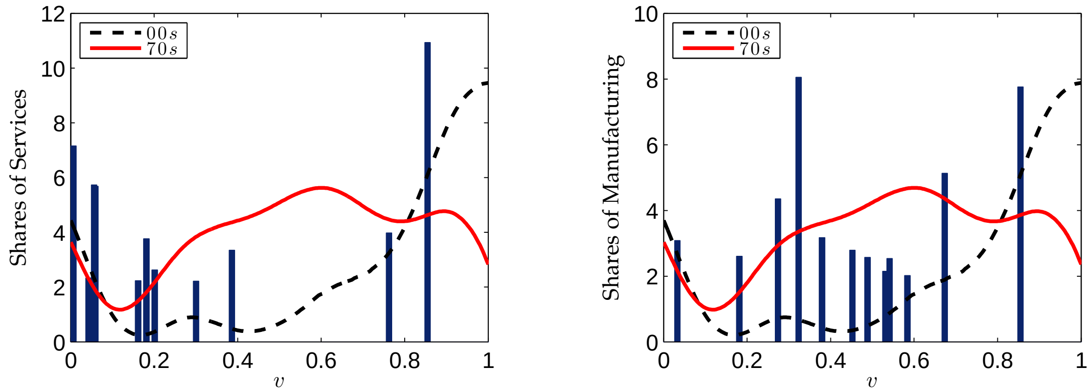
nificantly) for high v-occupations, but falls for intermediate ones. The picture is consistent with the phenomenon of *job polarization* as discussed in Section 2. This polarization feature is robust to different sample selection assumptions, see Appendix D for further details.

Figure 4 sharpens intuition concerning the relation of different v's to the data by showing the location of different sectoral occupations in the space of v's. The figure overlays the values of $b(v)$ with a bar graph displaying the employment shares of occupations belonging to particular sectors. Figure 4a does this for services and Figure 4b for manufacturing.[36] The service sector is associated mostly with extreme and, especially, "low" v occupations (the bar on the right in Figure 4a refers to managers and administrative support), while manufacturing is mostly middle v occupations (although with a wider range).

[34]In 2005 dollars we have $Y_{70} = \$36,998$ and $Y_{00} = \$45,260$. M is 302. In aggregate data using GDP deflator (table 1.1.9 in NIPA) and total non farm payroll (BLS) we get a value of real compensation per worker equal to $Y_{70} = \$37,114$ and $Y_{00} = \$53,304$. However deflating using CPI we get values consistent with our sample: $Y_{70} = \$37,966$ and $Y_{00} = \$45,151$.

[35]The LOWESS scatterplot smoothing builds up a smooth curve through a set of date points by fitting simple linear or quadratic models to localized subsets of data. We use a smoothing parameter of 0.4.

[36]Not shown are occupations that constitute less than 2% of the workforce of each sector.



(a) Services. (b) Manufacturing.

Figure 4: Occupations and v. Histograms: shares of occupations over v. Plots: smoothed values for $\log(b(v))$ over v and across decades.

5.3 ESTIMATING a

The envelope condition from the task choice component of the worker's equilibrium problem, $w(k) = \max_{v \in [\underline{v}, \bar{v}]} \omega(k, v)$, implies that:

$$\frac{d \log w(\tilde{k}(v))}{dk} = \frac{\partial \log \omega(\tilde{k}(v), v)}{\partial k} = \frac{\partial \log a(\tilde{k}(v), v)}{\partial k} = \frac{\partial \alpha}{\partial k}(k(v), v), \tag{19}$$

where $\alpha(k, v) := \log a(k, v)$. An empirical counterpart for $\frac{d \ln w(\tilde{k}(v))}{dk}$ is constructed in three steps. First, information from the CPS on weeks and usual hours worked in the previous year and self reported yearly labor income is used to impute workers' average hourly wages. Second, wages are averaged over occupation to construct empirical counterparts of $w(\tilde{k}(v))$. Third, a LOWESS smoother is applied to the log of this series and to $\tilde{k}$, derivatives of each function are calculated and $\frac{d \log w(\tilde{k}(v))}{dk} = \frac{d \log w(\tilde{k}(v))}{dv} \Big/ \frac{\partial \tilde{k}(v)}{\partial v}$ is found. Figure 5 displays $\log w(k)$ by talent across decades. From the 20^{th} to the 80^{th} talent percentile, log wages are linear in talent. In the 2000's, fast growth in wages occurs over the top two talent deciles, while in the 1970's it occurs only over the top decile. Finally for both decades (but especially for the 1970's) wages grow more rapidly over the bottom two deciles. In our benchmark quantitative environment we set:

$$\frac{\partial \alpha}{\partial k}(k, v) = \alpha_1 + \alpha_2 \cdot v. \tag{20}$$

Thus, α_2 captures comparative advantage. We later consider the alternative specification $\frac{\partial \alpha}{\partial k}(k, v) = \alpha_3 \cdot v^2$ in which comparative advantage is increasing with task complexity.

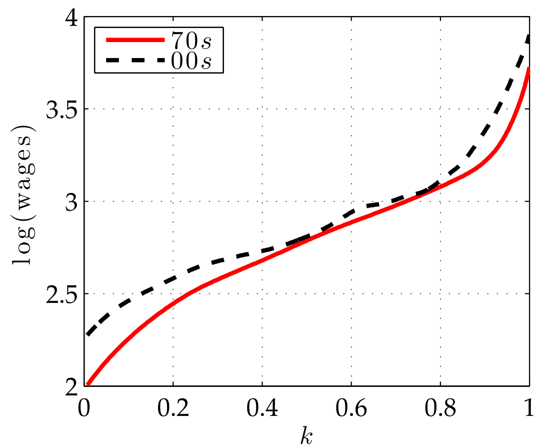


Figure 5: log(wages) over talents across decades.

Benchmark estimates of α_1 and α_2 are obtained by regressing $\frac{d \log w(\tilde{k}(v))}{dk}$ onto a constant and the task index v. The regression is weighted by the share of workers in each v. Results are reported in Table 2.[37] They show a significant increase in the comparative advantage parameter α_2 between the 1970's and 2000's. Loosely, this is driven by the increase in wage growth over high talents occurring between the 1970s and the 2000s.

Table 2: Estimation of Productivity Function.

Decade	**Parameters**	
	α_1	α_2
70s	1.07	1.72
	(0.25)	(0.28)
00s	0.41	3.01
	(0.32)	(0.22)

Notes: $N = 302$. Estimation of α_1 and α_2 from Equation (19). Standard errors in parenthesis.

To complete the parametrization of the production function it remains to determine TFP, A. This is obtained by dividing aggregate per capita income by the approximation

[37]Previous version of this paper estimated the value of α_2 looking at the distribution of the income distribution via a simulated method of moments. The identifying features being the variance and skewness of the distribution and it's changes over time. Using this procedure we also found a positive value of α_2 growing over time.

to the aggregator $\exp\{\int_{\underline{v}}^{\overline{v}} b(v) \log\{y(v)\} dv\}$:

$$A = \frac{Y}{\exp\{\sum_{m=1}^{M} b_m \log\{e^{(\alpha_1+\alpha_2 v_m)\tilde{k}_m} e_m S_m\}\}},$$

with e_m average hours worked in occupation m and α_1 and α_2 the previously estimated productivity parameters.

6 Quantitative Implications for Policy

In this section, we compute the optimal policy response to the technical change derived in Section 5. Calculation of policy requires a specification of worker and societal preferences and the amount of resources devoted to public spending. We briefly turn to this and then give our quantitative results.

6.1 Selection of Remaining Parameters and Computational Procedure

We assume that worker preferences are given by:

$$U(c,e) = \log c - \frac{e^{1+\gamma}}{1+\gamma}.$$

Note that the choice of U has no impact on the estimation of $b(v)$ and $a(k,v)$. We follow Chetty, Guren, Manoli, and Weber (2011) and set the Frisch labor supply elasticity to $1/\gamma = 0.75$.

We identify the share of output allocated to public spending with the aggregate tax to income ratio in our CPS sample. On this basis, $(G/Y)_{70} = 16.2\%$ and $(G/Y)_{00} = 14.0\%$; we set the G/Y ratio to the intermediate value of 15%.[38] Finally, in our benchmark calculations a utilitarian government is assumed: $g_k = \pi_k$ for all talents k.

To calculate optimal policy at our selected and estimated parameters, we first formulate the government's optimization as an optimal control problem. Details of this formulation are given in Appendix C. We then solve the problem numerically using the GPOPS-II software.[39]

[38]Using NIPA data (Table 1.1.6) would have implied that $(G/Y)_{70} = 23.9\%$ and $(G/Y)_{00} = 19.3\%$. However, since we are concerned with spending financed out of income taxation (paid by our subsample of labor income earners) we use the alternative CPS-generated estimates.

[39]GPOPS-II is a flexible software for solving optimal control problems. For additional details refer to

6.2 Optimal Tax Results

Table 3 reports optimal average and marginal tax rates as a function of income percentiles for the 1970s and the 2000s. Over this time period, average rates rise at low incomes and fall at high and, especially, middle incomes. Transfers to the lowest decile are reduced. Marginal rates fall at low and mid incomes and rise at higher incomes (except at those in the very top percentile, where marginal subsidies increase.)

Table 3: Optimal Tax Rates on Real Labor Income.

	Decade	**Percentiles of Income**					
		10th	*25th*	*50th*	*75th*	*90th*	*99th*
Averages	70s	-11.9	-7.3	6.9	22.3	26.1	22.2
	00s	-2.3	-1.1	5.6	19.9	26.1	21.9
Marginals	70s	20.3	34.1	44.3	40.3	23.9	-0.6
	00s	15.3	25.4	39.7	42.2	27.4	-2.2

To understand the evolution of optimal tax reported in Table 3, we return to the tax formula (2) derived earlier.

Accounting for Optimal Taxes

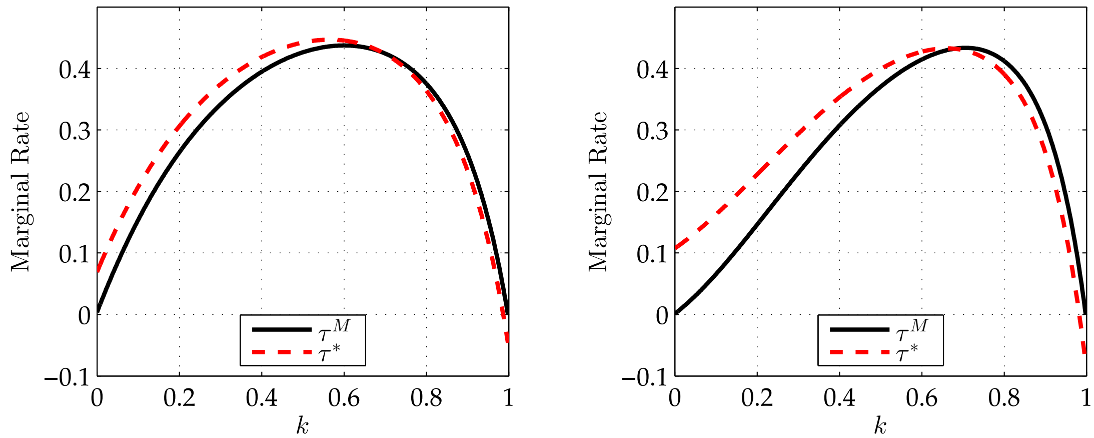
As observed previously, tax formula (2) allows us to decompose optimal tax rates into "Mirrleesian" and "Wage Compression" components. In particular, let τ_k^{M} denote the "Mirrleesian" marginal tax rate in the absence of the wage compression term:[40]

$$\tau_k^{\mathrm{M}} = \frac{\frac{\Delta w_{k+1}^*}{w_{k+1}^*}\frac{1-\Pi_k}{\pi_k}\mathcal{H}_k^*\Psi_k^*}{1+\frac{\Delta w_{k+1}^*}{w_{k+1}^*}\frac{1-\Pi_k}{\pi_k}\mathcal{H}_k^*\Psi_k^*},$$

and define the wage compression component to be the residual $\tau_k^{\mathrm{WC}} = \tau_k^* - \tau_k^{\mathrm{M}}$. In Figure 6, we plot the Mirrleesian component τ_k^M and the overall optimal marginal rate τ_k^* at each income percentile k and for each decade.

Patterson and Rao (2013).

[40]That is set the wage compression term to 0 in (2) and rearrange. For convenience, we continue to state tax formulas and their components in their discrete, rather than continuous forms.

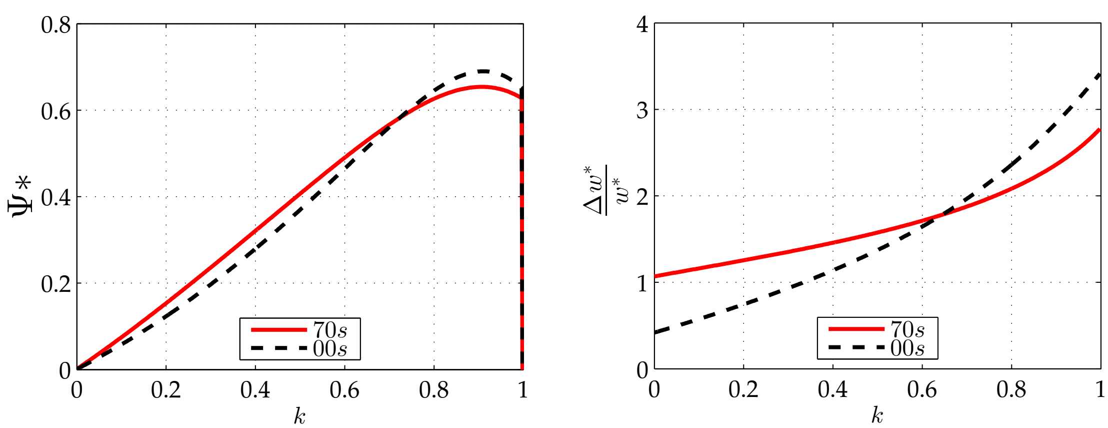


(a) 1970's. (b) 2000's.

Figure 6: Decomposing taxes. Solid curve: Mirrleesian tax, τ^M. Dashed curve: overall tax, τ^*.

Figure 6 shows that technical change deforms the Mirrleesian component pushing it to the right except at the lowest and highest talent. In addition, it raises the wage compression component at lower incomes and reduces it at higher ones. Overall the wage compression term becomes quantitatively more important.

EVOLUTION OF THE MIRRLEES TERM We further decompose the Mirrlees term into its redistributive Ψ and wage growth parts in Figure 7.[41]



(a) Ψ^*. (b) Wage Growth, $\frac{\Delta w^*}{w^*}$.

Figure 7: Mirrlees term decomposition.

[41]The other components are constant over time under our assumptions.

The main impact of technical change is upon wage growth (with some slight reinforcement from the redistributive term Ψ^*). This is largely driven by shifts to the a function. As noted previously, our estimates suggest that the productivities of low talents catch up with high in less complicated tasks and fall behind in more complex ones. At any effort profile and, in particular, at the optimal one, this shift compresses the wage distribution at the bottom and stretches it at the top. Shifts in the b function and in task demand from the middle to the extremes slightly reinforce the effect. The impact of the latter is, however, surprisingly small. This is largely because, in relevant areas of the task space, modest adjustments in the tasks of workers $\tilde{v}^*$ are consistent with quite large variations in the density of workers across tasks $\tilde{k}^*_v$. Consequently, increases in the demand for low and high tasks are met with increases in the number of workers performing these tasks, but relatively little adjustment in task assignment and, hence, relative productivities and wages, see Appendix E. The overall effect of these changes is to relax incentive constraints and reduce marginal taxes at the bottom, but to tighten them and raise marginal taxes at the top.

EVOLUTION OF THE WAGE COMPRESSION TERM Adjustment of the wage compression terms is in the opposite direction to the adjustment of the Mirrlees term previously described. Figure 8 displays this adjustment. It shows that the wage compression term *rises* at low incomes and *falls*, becoming more negative, at higher ones. These changes are

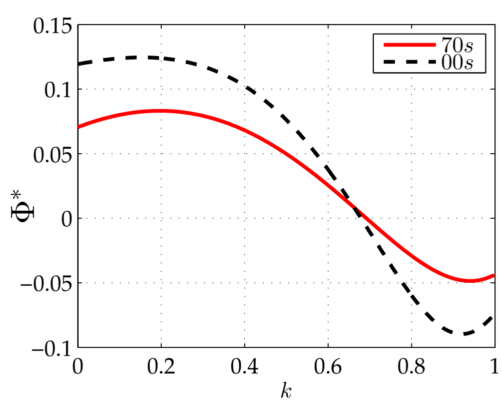
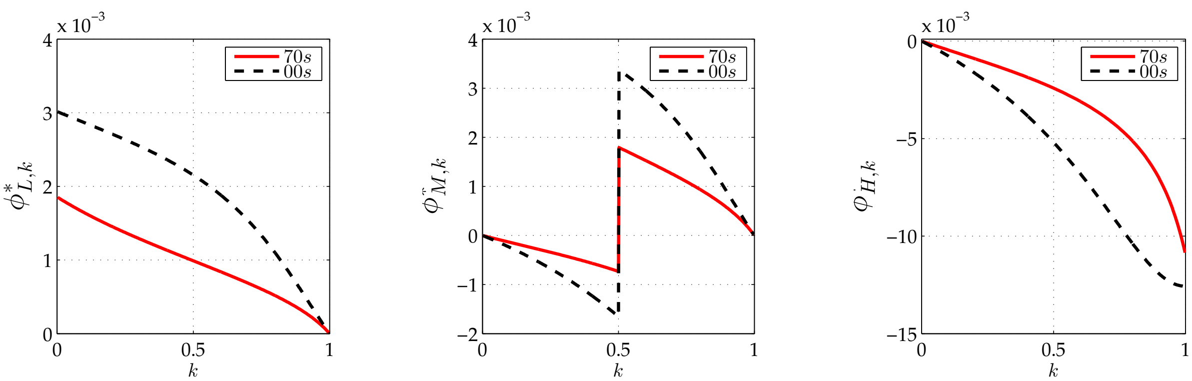


Figure 8: Evolution of the wage compression term.

largely attributable to adjustments in the relative wage elasticities $\phi^*_{k,j}$. The k-th talent's wage compression term is given by:

$$\Phi_k^* = \sum_{j=1}^{K-1} \mathcal{M}_{k,j} \phi_{k,j}^*. \tag{21}$$

Equation (21) expresses Φ_k^* as a weighted sum of relative wage elasticities, with the weights depending upon the marginal incentive benefit of adjusting each pair of relative wages. Mechanically, $\phi_{k,j}^*$ is positive if $j \geq k$ and negative otherwise, so that all $\phi_{k,j}^*$ are positive if $k = 1$ and all are negative if $k = K$. For some intermediate k, positive and negative terms cancel and the wage compression term is zero. An increase in the lowest talent's effort pushes all higher talents upwards through the task spectrum, raising the relative wages of all adjacent talents. This tightens all incentive constraints and is undesirable. Consequently, the lowest talent has the highest wage compression term and that talent's effort should be deterred at the margin. For the highest talent, this argument is reversed. An increase in the highest talent's effort pushes all lower talents downwards through the task set, compressing relative wages. This relaxes incentive constraints and should be encouraged at the margin with lower marginal taxes on high talent incomes. For intermediate talents these effects wholly or partially offset, leading to wage compression terms that are smaller in absolute value.



(a) e_L: 1970's and 2000's. (b) e_M: 1970's and 2000's. (c) e_H: 1970's and 2000's.

Figure 9: Effort elasticity of tasks.

Figure 9 shows the impact of technical change on relative wage elasticities (normalized by population shares) $\phi_{k,j}^*/\pi_k$, $j = 1, \ldots, K$ for low, mid and high talents (labelled L, M and H).[42] It indicates that almost all $\phi_{k,j}^*$ rise in absolute value. This is largely a

[42]Note the global impact of relative wages to an effort adjustment, an example of the ripple effect described previously.

consequence of the rise in the comparative adjustment parameter α_2 which, although it dampens the assignment response to adjustments in effort, raises the sensitivity of relative wages to any reassignment that occurs. Changes in the b function has only moderate effects on these elasticities, see Appendix E.

COMBINING TERMS Although the Mirrlees and wage compression terms evolve in opposite directions, it is the adjustment to the Mirrlees term that dominates over most incomes. Consequently, marginal tax rates fall at low and rise at high incomes (up to the top income percentiles), although not as much as they would have done absent adjustments to the wage compression term.

6.3 ALTERNATE PRODUCTIVITY FUNCTION

As discussed in Subsection 5.3, in our benchmark estimation the increase in the growth rate for log wages at higher talents leads to the identification of a growing comparative advantage over time. In this section we explore an alternative formulation of the productivity function aimed at fitting more closely the high and increasing growth rate of wages for high talents. Specifically, we set $\frac{\partial a}{\partial k}(k,v) = \alpha_3 \cdot v^2$. Proceeding as in Subsection 5.3 we find a value of $\alpha_3 = 0.79$ (0.03) for the 1970's and a value of $\alpha_3 = 0.91$ (0.04) for the 2000's.[43] As in our benchmark, there is an increase in the degree of comparative advantage over time. Given the quadratic nature of the productivity function the change in overall top to bottom talent inequality in wages is greater than in our benchmark setting. In Table 4 we display the resulting optimal behavior of average and marginal tax rates over percentiles of the income distribution.

Table 4: Optimal Tax Rates on Real Labor Income, Alternate Case.

	Decade	**Percentiles of Income**					
		10th	*25th*	*50th*	*75th*	*90th*	*99th*
Averages	70s	-3.3	0.9	7.8	20.3	24.7	20.8
	00s	-12.5	-8.7	4.0	21.6	28.3	23.7
Marginals	70s	17.2	28.8	40.0	38.9	24.3	-1.8
	00s	20.0	32.5	44.5	44.8	29.5	-1.5

Note: Estimates of tax rates determined using $\frac{\partial a}{\partial k} = \alpha_3 \cdot v^2$.

[43] In addition, to emphasize the behavior of wages for higher talents we estimate α_3 without weighting by the shares of talent in each occupation.

Relative to the benchmark, the striking difference is the sharp fall in average and increase in marginal rates over the time period.[44] Now, the rise in comparative advantage dominates. It strengthens the wage compression channel and increases wage growth across talents. This increases the motive for redistribution towards the bottom.

7 Conclusion

We relate the positive literature on technical change to normative work on optimal taxation by embedding an assignment model into an optimal tax framework. The assignment component induces an indirect production function over worker efforts enabling us to map technical parameters determining the productivity of task-talent matches and the demand for tasks to the variables and elasticities relevant for optimal tax analysis. We investigate the implications of changes in these parameters for optimal taxes, measure the extent of this change in US data and evaluate its implications for optimal policy.

The impacts of technical change on wage growth across talents and the substitutability of talents across tasks emerge as key drivers of policy. The twisting of the task-talent productivity function with low talents catching up in simple tasks and falling behind in more complex ones compresses wage differentials at the bottom, while expanding them at the top. It is a force for less redistribution and lower marginal taxes from the middle to the bottom and more redistribution and higher marginal taxes from the top to the middle. On the other hand, increased complementarity between talent and task complexity reduces the substitutability of talents and gives the government more tax leverage over the wage distribution. It is a force for higher marginal tax rates at the bottom. A key message of this paper is that policy depends upon the balance of these forces.

Theoretical public finance has traditionally focused on labor supply not demand and has rarely considered the implications of technological change for policy. It has much to learn from labor economics, but it is also well placed to develop the policy implications of labor economists' findings. Our paper is a first step in this direction. We conclude by describing three extensions that we leave for future research. First, our assignment model places strong restrictions on the distribution of workers across task productivities. These restrictions permit identification of key parameters and underpin our strategy for bringing the model to the data; they are also typical of a segment of the literature. However, they are inconsistent with intra-occupational wage variability except insofar as the latter reflects the coarseness of occupations as measures of tasks or errors in occupational coding. Second, the model assumes that the matching of talents to tasks is frictionless.

[44] This evolution is, however, much more similar to that of actual policy reported in Table 1.

Thus, our quantitative work is best viewed as capturing the long run policy response to technical change after the (possibly slow) reassignment of workers to tasks following such change.[45] The role of income taxation in supplementing other sources of insurance during transitions is omitted. Third, we have treated technical change parametrically. Relaxing these restrictions remain important topics for further research.[46]

References

Acemoğlu, D. (2002). Directed technical change. *The Review of Economic Studies 69*(4), 781–809.

Acemoğlu, D. and D. Autor (2011). Skills, tasks and technologies: Implications for employment and earnings. In D. Card and O. Ashenfelter (Eds.), *Handbook of Labor Economics*, Volume 4B, pp. 1043–1171. Elsevier Press.

Ales, L., M. Kurnaz, and C. Sleet (2013). Optimal taxation with publicly observable tasks. Carnegie Mellon University.

Autor, D. and D. Dorn (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review 103*(5), 1553–1597.

Autor, D., L. Katz, and M. Kearney (2006). The polarization of the us labor market. *American Economic Review 96*(2), 189–194.

Autor, D. H., L. F. Katz, and A. B. Krueger (1998). Computing inequality: Have computers changed the labor market? *Quarterly Journal of Economics 113*(4), 1169–1213.

Autor, D. H., F. Levy, and R. J. Murnane (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics 118*(4), 1279–1333.

Bonnans, J. and A. Shapiro (2000). *Perturbation Analysis of Optimization Problems*. Springer-Verlag: Springer Series in Operations Research.

Bresnahan, T. F., E. Brynjolfsson, and L. M. Hitt (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics 117*(1), 339–376.

Brito, D., J. Hamilton, S. Slutsky, and J. Stiglitz (1995). Randomization in optimal income tax schedules. *Journal of Public Economics 56*(2), 189–223.

[45]This long run perspective suggests an alternative application of our model in which effort is time devoted to converting talent into skill rather than time devoted to market work.

[46]The first is relaxed in Rothschild and Scheuer (2013), but in a setting with two tasks and no technical change. The third is partly relaxed in Slavík and Yazıcı (2014), but not in a setting with assignment.

Buera, F. and J. Kaboski (2012). The rise of the service economy. *American Economic Review 102*(6), 2540–2569.

Chetty, R., A. Guren, D. Manoli, and A. Weber (2011). Are micro and macro labor supply elasticities consistent? a review of evidence on the intensive and extensive margins. *The American Economic Review 101*(3), 471–475.

Costinot, A. and J. Vogel (2010). Matching and inequality in the world economy. *Journal of Political Economy 118*(4), 747–786.

Goldin, C. and L. F. Katz (1998). The origins of technology-skill complementarity. *The Quarterly Journal of Economics 113*(3), 693–732.

Goos, M. and A. Manning (2007). Lousy and lovely jobs: The rising polarization of work in britain. *Review of Economics and Statistics 89*(1), 118–133.

Goos, M., A. Manning, and A. Salomons (2009). Job polarization in europe. *American Economic Review: Papers & Proceedings 99*(2), 58–63.

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

Grossman, E. and E. Rossi-Hansberg (2008). Trading tasks: A simple theory of offshoring. *American Economic Review 98*(5), 1978–1997.

Heathcote, J., F. Perri, and G. L. Violante (2010). Unequal we stand: An empirical analysis of economic inequality in the united states, 1967–2006. *Review of Economic Dynamics 13*(1), 15–51.

Heathcote, J., K. Storesletten, and G. Violante (2014). Optimal tax progressivity: An analytical framework. Staff Report 496. FRB Minneapolis.

Hellwig, M. (2007). The undesirability of randomized income taxation under decreasing risk aversion. *Journal of Public Economics 91*(3-4), 791–806.

Krusell, P., L. Ohanian, J. Ríos-Rull, and G. Violante (2000). Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica 68*(5), 1029–1053.

Milgrom, P. and C. Shannon (1994). Monotone comparative statics. *Econometrica 62*(1), 157–180.

Mirrlees, J. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies 38*(2), 175–208.

Mokyr, J. (1992). *The lever of riches: Technological creativity and economic progress*. Oxford University Press, USA.

Naito, H. (1999). Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency. *Journal of Public Economics 71*(2), 165–188.

Patterson, M. A. and A. V. Rao (2013). GPOPS-II: A MATLAB software for solving multiple-phase optimal control problems using hp–adaptive gaussian quadrature collocation methods and sparse nonlinear programming. Working Paper.

Rothschild, C. and F. Scheuer (2013). Redistributive taxation in the roy model. *Quarterly Journal of Economics 128*(2), 623–662.

Roy, A. (1950). The distribution of earnings and of individual output. *Economic Journal 60*(239), 489–501.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *Review of Economic Studies 68*(1), 205–229.

Salanié, B. (2011). *Economics of Taxation*. MIT Press, USA.

Sattinger, M. (1975). Comparative advantage and the distributions of earnings and abilities. *Econometrica 43*(3), 455–468.

Slavík, C. and H. Yazıcı (2014). Machines, buildings and optimal dynamic taxes. Forthcoming, *Journal of Monetary Economics*.

Stiglitz, J. (1982). Self selection and pareto efficient taxation. *Journal of Public Economics 17*(2), 213–240.

Teulings, C. (1995). The wage distribution in a model of the assignment of skills to jobs. *Journal of Political Economy 103*(2), 280–315.

Appendix: For Online Publication

A Proofs from Section 3

We first formalize a planning problem in a space of simple direct mechanisms and then recover an optimal tax equilibrium from its solution.

Implementation via simple, direct mechanisms A simple, direct mechanism is a tuple $\{c_k, e_k\}_{k=1}^K$ of message-contingent consumptions and effort recommendations and a severe penalty. The mechanism implicitly defines a set of shadow wages and incomes: $w_k = F_k(\pi_1 e_1, \ldots, \pi_K e_K)$ and $q_k = w_k e_K$. The planner and the workers play the following game. First, the planner selects a mechanism. Second, each worker sends a message $k \in \{1, \ldots, K\}$, exerts an effort and generates a publicly observable income q. Third, if a worker sends the message k and generates the income q_k, she receives the consumption c_k. Otherwise she receives the severe penalty. The latter is assumed to be sufficient to deter a worker who has sent message k from generating an income $q \neq q_k$. Without loss of generality, attention is restricted to mechanisms that induce workers to report truthfully conditional on almost all other workers doing so. Thus, the mechanism is constrained to satisfy, for all $k \in \{1, \ldots, K\}$ and $j \in \{1, \ldots, K\} \backslash \{k\}$,

$$U(c_k, e_k) \geq U(c_j, q_j / w_k). \tag{A.1}$$

The planner's problem is then:

$$\sup \sum_{k=1}^{K} U(c_k, e_k) g_k \tag{A.2}$$

s.t. for all $k \in \{1, \ldots, K\}$ and $j \in \{1, \ldots, K\} \backslash \{k\}$

$$U(c_k, e_k) \geq U\left(c_j, \frac{F_j(e_1\pi_1, \ldots, e_K\pi_K)}{F_k(e_1\pi_1, \ldots, e_K\pi_K)} e_j\right). \tag{A.3}$$

and

$$G + \sum_{k=1}^{K} c_k \pi_k \leq F(e_1\pi_1, \ldots, e_K\pi_K). \tag{A.4}$$

Let $\{c_k^*, e_k^*\}_{k=1}^K$ denote a solution to (A.2) with corresponding shadow wages $\{w_k^*\}_{k=1}^K$, $w_k^* = F_k(e_1^*\pi_1, \ldots, e_K^*\pi_K)$. Agent types may be ordered according to their shadow wages and relabeled accordingly.[47] Consequently, there is no loss of generality in assuming that a k-th type has the k-th highest wage; we make this assumption below. A (k, j)-th incentive compatibility constraint (A.3) is local if $j \in \{k-1, k+1\}$. Lemma A.1 below is a well known consequence of the Spence-Mirrlees single crossing property and the structure of the incentive constraints in settings with exogenous wages; it continues to hold in the present setting.[48]

Lemma A.1. *Assuming types are labeled according to their ranking in the wage distribution, then (i) $c_{k+1} \geq c_k$ and $q_{k+1} \geq q_k$, (ii) only "local" incentive constraints bind.*

Proof of Proposition 1. Let χ^* denote the optimal multiplier on the resource constraint and $\eta_{k,j}^*$ the optimal multiplier on the (k, j)-th incentive constraint. In light of the previous lemma only local incentive constraints are potentially binding and, hence, only the $\eta_{k,k-1}^*$ and $\eta_{k,k+1}^*$ are potentially non-zero. The first order condition for e_k^* reduces to:

$$-U_e(c_k^*, e_k^*) = \frac{\chi^* w_k^* \pi_k}{D_k^*},$$

[47] Formally, let $\kappa : \{1, \ldots, K\} \to \{1, \ldots, K\}$ denote a permutation of $1, \ldots, K$ such that for $k = 1, \ldots, K-1$, $w_{\kappa(k+1)}^* \geq w_{\kappa(k)}^*$. Then types may be relabeled according to $k' = \kappa^{-1}(k)$.

[48] We omit the proof. It can be shown as an application of Theorems 3 and 4 in Milgrom and Shannon (1994).

where:

$$D_k^* := g_k + \eta_{k,k-1}^* - \eta_{k+1,k}^* \frac{U_e\left(c_k^*, \frac{q_k^*}{w_{k+1}^*}\right)}{U_e(c_k^*, e_k^*)} \frac{w_k^*}{w_{k+1}^*} + \eta_{k,k+1}^* - \eta_{k-1,k}^* \frac{U_e\left(c_k^*, \frac{q_k^*}{w_{k-1}^*}\right)}{U_e(c_k^*, e_k^*)} \frac{w_k^*}{w_{k-1}^*} + \frac{\chi^* \pi_k}{U_c(c_k^*, e_k^*)} \Phi_k^* + \frac{\chi^* \pi_k}{U_c(c_k^*, e_k^*)} \Upsilon_k^*,$$

$$\Phi_k^* := \frac{U_c(c_k^*, e_k^*)}{\pi_k} \sum_{j=1}^{K-1} \frac{\eta_{j+1,j}^*}{\chi^*} \frac{U_e\left(c_j^*, \frac{q_j^*}{w_{j+1}^*}\right)}{U_e(c_k^*, e_k^*)} \frac{w_j^* e_j^*}{w_{j+1}^* e_k^*} \phi_{k,j}^*,$$

and

$$\Upsilon_k^* := \frac{U_c(c_k^*, e_k^*)}{\pi_k} \sum_{j=1}^{K-1} \frac{\eta_{j-1,j}^*}{\chi^*} \frac{U_e\left(c_j^*, \frac{q_j^*}{w_{j-1}^*}\right)}{U_e(c_k^*, e_k^*)} \frac{w_j^* e_j^*}{w_{j+1}^* e_k^*} \phi_{k,j-1}^*.$$

The first order condition for c_k^* reduces to:

$$U_c(c_k^*, e_k^*) = \frac{\chi^* \pi_k}{g_k + \eta_{k,k-1}^* - \eta_{k+1,k}^* \frac{U_c\left(c_k^*, \frac{q_k^*}{w_{k+1}^*}\right)}{U_c(c_k^*, e_k^*)} + \eta_{k,k+1}^* - \eta_{k-1,k}^* \frac{U_c\left(c_k^*, \frac{q_k^*}{w_{k-1}^*}\right)}{U_c(c_k^*, e_k^*)}}.$$

Define the consumption-effort wedge:

$$\frac{\tau_k^*}{1-\tau_k^*} = -\frac{w_k^* U_c(c_k^*, e_k^*)}{U_e(c_k^*, e_k^*)} - 1.$$

Combining expressions gives:

$$\frac{\tau_k^*}{1-\tau_k^*} = \frac{U_c(c_k^*, e_k^*)}{\pi_k} \left\{ \frac{\eta_{k+1,k}^*}{\chi^*} \left\{ \frac{U_c\left(c_k^*, \frac{q_k^*}{w_{k+1}^*}\right)}{U_c(c_k^*, e_k^*)} - \frac{U_e\left(c_k^*, \frac{q_k^*}{w_{k+1}^*}\right)}{U_e(c_k^*, e_k^*)} \frac{w_k^*}{w_{k+1}^*} \right\} + \frac{\eta_{k-1,k}^*}{\chi^*} \left\{ \frac{U_c\left(c_k^*, \frac{q_k^*}{w_{k-1}^*}\right)}{U_c(c_k^*, e_k^*)} - \frac{U_e\left(c_k^*, \frac{q_k^*}{w_{k-1}^*}\right)}{U_e(c_k^*, e_k^*)} \frac{w_k^*}{w_{k-1}^*} \right\} \right\} + \Phi_k^* + \Upsilon_k^*.$$

Under (NUIC), for all k, $\eta_{k,k+1}^* = 0$, $\Upsilon_k^* = 0$ and the previous expression reduces to:

$$\frac{\tau_k^*}{1-\tau_k^*} = \frac{U_c(c_k^*, e_k^*)}{\pi_k} \left\{ \frac{\eta_{k+1,k}^*}{\chi^*} \left\{ \frac{U_c\left(c_k^*, \frac{q_k^*}{w_{k+1}^*}\right)}{U_c(c_k^*, e_k^*)} - \frac{U_e\left(c_k^*, \frac{q_k^*}{w_{k+1}^*}\right)}{U_e(c_k^*, e_k^*)} \frac{w_k^*}{w_{k+1}^*} \right\} \right\} + \Phi_k^*.$$

Again under (NUIC), the first order equation for c_{k+1}^* implies the following recursion for

$\eta^*_{k+1,k}$:

$$\frac{\eta^*_{k+1,k}}{\chi^*} = \left(1 - \frac{g_{k+1}U_c(c^*_{k+1}, e^*_{k+1})}{\chi^* \pi_{k+1}}\right) \frac{\pi_{k+1}}{U_c(c^*_{k+1}, e^*_{k+1})} + \frac{\eta^*_{k+2,k+1}}{\chi^*} \frac{U_c\left(c^*_{k+1}, \frac{q^*_{k+1}}{w^*_{k+2}}\right)}{U_c(c^*_{k+1}, e^*_{k+1})},$$

with $\eta_{K^*+1,K^*} = 0$. Iterating on this recursion, setting $\Psi^*_k := \frac{\eta^*_{k+1,k}}{\chi^*} \frac{U_c(c^*_k, e^*_k)}{1-\Pi_k}$ and using the definition of $\mathcal{H}^*_k$ in the proposition gives (2). A standard appeal to the taxation principle establishes that the solution to the planning problem is an optimal tax equilibrium allocation and that τ^*_k is the optimal marginal tax rate for the k-th type. □

B PROOFS FROM SECTION 4

Proof of Proposition 2. Let T, $\{c_k, e_k, \lambda_k\}_{k=1}^K$ and $\{\omega_k\}_{k=1}^K$ denote a tax equilibrium at spending level G. Since workers of a given type select the highest possible wage, it follows that for each k there is a $w_k < \infty$ such that for every $v \in \text{Supp}\, \Lambda_k$, $\omega_k(v) = w_k$ and for $v \notin \text{Supp}\, \Lambda_k$, $\omega_k(v) \leq w_k$. Without loss of generality, assume that the firm's first order conditions hold at each k and almost every $v \in \Lambda_k(v)$:

$$\omega_k(v) = b(v) \left(\frac{Y}{y(v)}\right)^{\frac{1}{\varepsilon}} a_k(v).$$

If $v \in [\underline{v}, \overline{v}] \backslash \cup_{k=1}^k \text{Supp}(\Lambda_k)$, then $y(v) = 0$ and for all k, $\omega_k(v) = b(v) \left(\frac{Y}{y(v)}\right)^{\frac{1}{\varepsilon}} a_k(v) = \infty > w_k$. Since this a contradiction, $[\underline{v}, \overline{v}] \backslash \cup_{k=1}^k \text{Supp}(\Lambda_k)$ must be of measure zero and almost all tasks are performed. Without loss of generality, we select versions of tax equilibria in which all tasks are performed. For all v and v' in Supp Λ_k with $v > v'$,

$$1 = \frac{\omega_k(v)}{\omega_k(v')} = \frac{b(v) \left(\frac{Y}{y(v)}\right)^{\frac{1}{\varepsilon}} a_k(v)}{b(v') \left(\frac{Y}{y(v')}\right)^{\frac{1}{\varepsilon}} a_k(v')} < \frac{b(v) \left(\frac{Y}{y(v)}\right)^{\frac{1}{\varepsilon}} a_{k+j}(v)}{b(v') \left(\frac{Y}{y(v')}\right)^{\frac{1}{\varepsilon}} a_{k+j}(v')} < \frac{\omega_{k+j}(v)}{\omega_{k+j}(v')}.$$

It follows that $v' \notin \Lambda_{k+j}$ and so $\sup \Lambda_k \leq \inf \Lambda_{k+j}$. Since the supports Λ_k cover $[\underline{v}, \overline{v}]$, it follows that they partition $[\underline{v}, \overline{v}]$ into sub-intervals $[\tilde{v}_0, \tilde{v}_1]$, $[\tilde{v}_1, \tilde{v}_2], \ldots, \tilde{v}_{K-1}, \tilde{v}_K]$, with $\tilde{v}_0 = \underline{v}$, $\tilde{v}_K = \overline{v}$ and cl $\Lambda_k = [\tilde{v}_{k-1}, \tilde{v}_k]$. By assumption each Λ_k has a density λ_k (concentrated on $[\tilde{v}_{k-1}, \tilde{v}_k]$). Since $w_k = \omega_k(v)$, $v \in (\tilde{v}_{k-1}, \tilde{v}_k)$, we have for all such v:

$$w_k = b(v) \left(\frac{Y}{a_k(v) e_k \lambda_k(v)}\right)^{\frac{1}{\varepsilon}} a_k(v). \tag{B.1}$$

Hence, from the labor market clearing condition:

$$\pi_k = \int_{\tilde{v}_{k-1}}^{\tilde{v}_k} \lambda_k(v)dv = \frac{Y}{w_k^\varepsilon e_k} \int_{v_{k-1}}^{v_k} b(v)^\varepsilon a_k(v)^{\varepsilon-1}$$

And so:

$$w_k = B_k(\tilde{v}_{k-1}, \tilde{v}_k) \left(\frac{Y}{\pi_k e_k}\right)^{\frac{1}{\varepsilon}}, \tag{B.2}$$

where $B_k(\tilde{v}_{k-1}, \tilde{v}_k) := \left[\int_{\underline{v}_{k-1}}^{\overline{v}_k} b(v)^\varepsilon a_k(v)^{\varepsilon-1} dv\right]^{\frac{1}{\varepsilon}}$. Substituting (B.1) into (B.2) gives for $v \in (\tilde{v}_{k-1}, \tilde{v}_k)$,

$$\lambda_k(v) = \frac{b(v)^\varepsilon a_k(v)^{\varepsilon-1}}{B_k(\tilde{v}_{k-1}, \tilde{v}_k)^\varepsilon} \pi_k.$$

In addition, for $v < \tilde{v}_{k-1}$ and $v > \tilde{v}_k$, $\lambda_k(v) = 0$.

Now for $v \in (\tilde{v}_{k-1}, \tilde{v}_k)$, $w_{k+1} > \omega_{k+1}(v) = b(v)\left(\frac{Y}{y(v)}\right)^{\frac{1}{\varepsilon}} a_{k+1}(v)$ and $w_k = \omega_k(v) = b(v)\left(\frac{Y}{y(v)}\right)^{\frac{1}{\varepsilon}} a_k(v)$. Hence: $\frac{w_{k+1}}{w_k} > \frac{a_{k+1}(v)}{a_k(v)}$. Conversely, for $v \in (\tilde{v}_k, \tilde{v}_{k+1})$, $w_{k+1} = \omega_{k+1}(v) = b(v)\left(\frac{Y}{y(v)}\right)^{\frac{1}{\varepsilon}} a_{k+1}(v)$ and $w_k > \omega_k(v) = b(v)\left(\frac{Y}{y(v)}\right)^{\frac{1}{\varepsilon}} a_k(v)$. Consequently, $\frac{w_{k+1}}{w_k} < \frac{a_{k+1}(v)}{a_k(v)}$. Then, by continuity of a_k and a_{k+1}, $\frac{w_{k+1}}{w_k} = \frac{a_{k+1}(\tilde{v}_k)}{a_k(\tilde{v}_k)}$. Combining the last equality with (B.2) gives the desired expression in the proposition.

Finally, given the effort allocation $\{e_k\}_{k=1}^K$ consider assigning workers so as to maximize output, i.e. solving:

$$\max_{\{\lambda_k\}} \left[\int_{\underline{v}}^{\overline{v}} b(v) \left\{\lambda_k(v) e_k a_k(v)\right\}^{\frac{\varepsilon-1}{\varepsilon}} dv\right]^{\frac{\varepsilon}{\varepsilon-1}}.$$

subject to for each k, $\pi_k = \int_{\underline{v}}^{\overline{v}} \lambda_k(v)dv$. This is a strictly concave maximization whose unique solution is determined by the first order conditions. Straightforward manipulation of these conditions establishes that the λ_k solved for above attains the solution to this problem. □

Proof of Lemma 1. Totally differentiating: $\frac{a_{j+1}}{a_j}(\tilde{v}_j) = \frac{B_{j+1}(\tilde{v}_j, \tilde{v}_{j+1})}{B_j(\tilde{v}_{j-1}, \tilde{v}_j)}\left(\frac{e_j \pi_j}{e_{j+1}\pi_{j+1}}\right)^{\frac{1}{\varepsilon}}$, with respect to $\tilde{v}_{j-1}$ and $\tilde{v}_j$ holding e_j/e_{j+1} fixed gives:

$$\frac{\partial \log \tilde{v}_j}{\partial \log \tilde{v}_{j-1}} = \frac{-\frac{\partial \log B_j}{\partial \log \tilde{v}_{j-1}}}{\frac{\partial \log a_{j+1}/a_j}{\partial \log \tilde{v}_j} - \frac{\partial \log B_{j+1}}{\partial \log \tilde{v}_j} + \frac{\partial \log B_j}{\partial \log \tilde{v}_j} - \frac{\partial \log B_{j+1}}{\partial \log \tilde{v}_{j+1}} \frac{\partial \log \tilde{v}_{j+1}}{\partial \log \tilde{v}_j}}.$$

Let $L_{j-1} = -\frac{\partial \log B_j}{\partial \log \tilde{v}_{j-1}} - \frac{\partial \log B_j}{\partial \log \tilde{v}_j}\frac{\partial \log \tilde{v}_j}{\partial \log \tilde{v}_{j-1}}$. It follows that:

$$L_{j-1} = -\frac{\partial \log B_j}{\partial \log \tilde{v}_{j-1}} \left\{ 1 - \frac{\frac{\partial \log B_j}{\partial \log \tilde{v}_j}}{\frac{\partial \log a_{j+1}/a_j}{\partial \log \tilde{v}_j} + \frac{\partial \log B_j}{\partial \log \tilde{v}_j} + L_j} \right\}.$$

Thus, if $L_j > 0$, then $L_{j-1} > 0$. For $j = K-1$, $\tilde{v}_{j+1} = \tilde{v}_K = \overline{v}$ and $\frac{\partial \log \tilde{v}_{j+1}}{\partial \log \tilde{v}_j} = \frac{\partial \log \tilde{v}_K}{\partial \log \tilde{v}_{K-1}} = 0$. Hence, $L_{K-1} = -\frac{\partial \log B_K}{\partial \log \tilde{v}_{K-1}} > 0$. It follows by induction that for all $j \in \{k, \ldots, K-2\}$, $L_j > 0$ and, hence,

$$\frac{\partial \log \tilde{v}_j}{\partial \log \tilde{v}_{j-1}} = \frac{-\frac{\partial \log B_j}{\partial \log \tilde{v}_{j-1}}}{\frac{\partial \log a_{j+1}/a_j}{\partial \log \tilde{v}_j} + \frac{\partial \log B_j}{\partial \log \tilde{v}_j} + L_j} > 0.$$

Similarly, for all $j \in \{1, \ldots, k-1\}$,

$$\frac{\partial \log \tilde{v}_j}{\partial \log \tilde{v}_{j+1}} = \frac{\frac{\partial \log B_{j+1}}{\partial \log \tilde{v}_{j+1}}}{\frac{\partial \log a_{j+1}/a_j}{\partial \log \tilde{v}_j} - \frac{\partial \log B_{j+1}}{\partial \log \tilde{v}_j} + \frac{\partial \log B_j}{\partial \log \tilde{v}_j} - \frac{\partial \log B_j}{\partial \log \tilde{v}_{j-1}}\frac{\partial \log \tilde{v}_{j-1}}{\partial \log \tilde{v}_j}}.$$

Let $M_{j+1} = \frac{\partial \log B_{j+1}}{\partial \log \tilde{v}_{j+1}} - \frac{\partial \log B_{j+1}}{\partial \log \tilde{v}_j}\frac{\partial \log \tilde{v}_j}{\partial \log \tilde{v}_{j+1}}$. It follows that:

$$M_{j+1} = \frac{\partial \log B_{j+1}}{\partial \log \tilde{v}_{j+1}} \left\{ 1 - \frac{\frac{\partial \log B_{j+1}}{\partial \log \tilde{v}_j}}{\frac{\partial \log a_{j+1}/a_j}{\partial \log \tilde{v}_j} - \frac{\partial \log B_{j+1}}{\partial \log \tilde{v}_j} + M_j} \right\}.$$

Thus, if $M_j > 0$, then $M_{j+1} > 0$. For $j = 1$, $\tilde{v}_{j-1} = \tilde{v}_0 = \underline{v}$ and $\frac{\partial \log \tilde{v}_{j-1}}{\partial \log \tilde{v}_j} = \frac{\partial \log \tilde{v}_0}{\partial \log \tilde{v}_1} = 0$. Hence, $M_1 = \frac{\partial \log B_1}{\partial \log \tilde{v}_1} > 0$. It follows by induction that for all $j \in \{1, \ldots, k-1\}$, $M_j > 0$ and, hence,

$$\frac{\partial \log \tilde{v}_j}{\partial \log \tilde{v}_{j+1}} = \frac{\frac{\partial \log B_{j+1}}{\partial \log \tilde{v}_{j+1}}}{\frac{\partial \log a_{j+1}/a_j}{\partial \log \tilde{v}_j} - \frac{\partial \log B_{j+1}}{\partial \log \tilde{v}_j} + M_j} > 0.$$

Next, taking logs and totally differentiating $\frac{a_{k+1}}{a_k}(\tilde{v}_k) = \frac{B_{k+1}(\tilde{v}_k, \tilde{v}_{k+1})}{B_j(\tilde{v}_{k-1}, \tilde{v}_k)} \left(\frac{e_k \pi_k}{e_{k+1}\pi_{k+1}}\right)^{\frac{1}{\varepsilon}}$ with respect to $\log e_k$ gives:

$$\frac{\partial \log \tilde{v}_k}{\partial \log e_k} = \frac{1}{\varepsilon} \left\{ \frac{1}{\frac{\partial \log a_{k+1}/a_k}{\partial \log \tilde{v}_k} + M_k + L_k} \right\} > 0.$$

Similarly, taking logs and totally differentiating $\frac{a_k}{a_{k-1}}(\tilde{v}_{k-1}) = \frac{B_k(\tilde{v}_{k-1},\tilde{v}_k)}{B_j(\tilde{v}_{k-2},\tilde{v}_{k-1})}\left(\frac{e_{k-1}\pi_{k-1}}{e_k\pi_k}\right)^{\frac{1}{\varepsilon}}$ with respect to $\log e_k$ gives:

$$\frac{\partial \log \tilde{v}_{k-1}}{\partial \log e_k} = -\frac{1}{\varepsilon}\left\{\frac{1}{\frac{\partial \log a_k/a_{k-1}}{\partial \log \tilde{v}_{k-1}} + M_{k-1} + L_{k-1}}\right\} < 0.$$

The implications for the elasticities $\phi_{k,j}$ described in the lemma then follow immediately from (12). □

Complete characterization of the sensitivity of task thresholds to the perturbation of a given talent's effort is provided in the next lemma.

Lemma B.1. *The threshold sensitivities satisfy:*

$$\frac{\partial \log \tilde{v}_j}{\partial \log e_k} = (\delta_{j,k-1} - \delta_{j,k})\frac{1}{\varepsilon},$$

where:

$$\delta_{j,k} = \begin{cases} (-1)^{j+k}\prod_{l=j}^{k-1}\left(-\frac{\tilde{v}_l}{B_l}\frac{\partial B_l}{\partial \tilde{v}_l}\right) n_{j-1}m_{k+1}/n_{K-1} & 1 \leq j \leq k \leq K-1 \\ (-1)^{j+k}\prod_{l=k}^{j-1}\left(\frac{\tilde{v}_l}{B_{l+1}}\frac{\partial B_{l+1}}{\partial \tilde{v}_l}\right) n_{k-1}m_{j+1}/n_{K-1} & K-1 \geq j > k \geq 1, \end{cases}$$

for $j = 1, \ldots, K-1$, $\delta_{j,K} = \delta_{j,0} = 0$, the n_i satisfy the recursion, $i = 2, \ldots, K-1$,

$$n_i = \left\{\frac{\tilde{v}_i}{a_{i+1}/a_i}\frac{\partial(a_{i+1}/a_i)}{\partial \tilde{v}_i} - \frac{\tilde{v}_i}{B_{i+1}}\frac{\partial B_{i+1}}{\partial \tilde{v}_i} + \frac{\tilde{v}_i}{B_i}\frac{\partial B_i}{\partial \tilde{v}_i}\right\} n_{i-1} + \left(\frac{\tilde{v}_i}{B_i}\frac{\partial B_i}{\partial \tilde{v}_i}\right)\left(\frac{\tilde{v}_{i-1}}{B_i}\frac{\partial B_i}{\partial \tilde{v}_{i-1}}\right) n_{i-2}$$

with $n_0 = 1$ and $n_1 = \frac{\tilde{v}_1}{a_2/a_1}\frac{\partial(a_2/a_1)}{\partial \tilde{v}_1} - \frac{\tilde{v}_1}{B_2}\frac{\partial B_2}{\partial \tilde{v}_1} + \frac{\tilde{v}_1}{B_1}\frac{\partial B_1}{\partial \tilde{v}_1}$ and the m_i satisfy the recursion, $i = K-2, \ldots, 1,$:

$$m_i = \left\{\frac{\tilde{v}_i}{a_{i+1}/a_i}\frac{\partial(a_{i+1}/a_i)}{\partial \tilde{v}_i} - \frac{\tilde{v}_i}{B_{i+1}}\frac{\partial B_{i+1}}{\partial \tilde{v}_i} + \frac{\tilde{v}_i}{B_i}\frac{\partial B_i}{\partial \tilde{v}_i}\right\} m_{i+1} + \left(\frac{\tilde{v}_{i+1}}{B_{i+1}}\frac{\partial B_{i+1}}{\partial \tilde{v}_{i+1}}\right)\left(\frac{\tilde{v}_i}{B_{i+1}}\frac{\partial B_{i+1}}{\partial \tilde{v}_i}\right) m_{i+2}.$$

with $m_{K-1} = \frac{\tilde{v}_{K-1}}{a_K/a_{K-1}}\frac{\partial(a_K/a_{K-1})}{\partial \tilde{v}_{K-1}} - \frac{\tilde{v}_{K-1}}{B_K}\frac{\partial B_K}{\partial \tilde{v}_{K-1}} + \frac{\tilde{v}_{K-1}}{B_{K-1}}\frac{\partial B_{K-1}}{\partial \tilde{v}_{K-1}}$ and $m_K = 1$.

Proof of Lemma B.1. Given an (equilibrium) effort profile $\{e_k\}_{k=1}^K$, (equilibrium) production maximizing task thresholds $\{\tilde{v}_k\}$ are determined by the conditions, $k = 1, \ldots, K-1$,

$$\frac{B_k(\tilde{v}_{k-1}, \tilde{v}_k)}{\{\pi_k e_k\}^{\frac{1}{\varepsilon}}} = \frac{B_{k+1}(\tilde{v}_k, \tilde{v}_{k+1})}{\{\pi_{k+1}e_{k+1}\}^{\frac{1}{\varepsilon}}}\frac{a_k(\tilde{v}_k)}{a_{k+1}(\tilde{v}_k)}.$$

with $\tilde{v}_0 = \underline{v}$ and $\tilde{v}_K = \overline{v}$. Hence, there are $K-1$ unknowns (and $K-1$ equations). The threshold sensitivities may be computed by taking logs in the preceding equations and totally differentiating with respect to $\log e_k$. This leads to the equations:

$$\Gamma \Delta v_k = E_k,$$

where:

$$\Gamma := \begin{pmatrix} \alpha_1 \tilde{v}_1 - \frac{\tilde{v}_1}{B_2}\frac{\partial B_2}{\partial \tilde{v}_1} + \frac{\tilde{v}_1}{B_1}\frac{\partial B_1}{\partial \tilde{v}_1} & -\frac{\tilde{v}_2}{B_2}\frac{\partial B_2}{\partial \tilde{v}_2} & 0 & 0 & \dots & 0 \\ \frac{\tilde{v}_1}{B_2}\frac{\partial B_2}{\partial \tilde{v}_1} & \alpha_2 \tilde{v}_2 - \frac{\tilde{v}_2}{B_3}\frac{\partial B_3}{\partial \tilde{v}_2} + \frac{\tilde{v}_2}{B_2}\frac{\partial B_2}{\partial \tilde{v}_2} & -\frac{\tilde{v}_3}{B_3}\frac{\partial B_3}{\partial \tilde{v}_3} & 0 & \dots & 0 \\ 0 & \frac{\tilde{v}_2}{B_3}\frac{\partial B_3}{\partial \tilde{v}_2} & \alpha_3 \tilde{v}_3 - \frac{\tilde{v}_3}{B_4}\frac{\partial B_4}{\partial \tilde{v}_3} - \frac{\tilde{v}_3}{B_3}\frac{\partial B_3}{\partial \tilde{v}_3} & -\frac{\tilde{v}_4}{B_4}\frac{\partial B_4}{\partial \tilde{v}_4} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

$\Delta v_k = (\frac{\partial \log v_1}{\partial \log e_k} \dots \frac{\partial \log v_k}{\partial \log e_k} \dots \frac{\partial \log v_{K-1}}{\partial \log e_k})'$ and $E_k = (0 \dots - \frac{1}{\varepsilon}\ \frac{1}{\varepsilon} \dots 0)'$ with non-zero elements in the $k-1$ and k-th rows. Thus,

$$\Delta v_k = \Gamma_k^{-1} E_k.$$

and in fact:

$$\frac{\partial \log v_j}{\partial \log e_k} = (\delta_{j,k-1} - \delta_{j,k})\frac{1}{\varepsilon},$$

where $\delta_{j,k}$ is the (j,k)-th element of Γ_k^{-1}. Since Γ_k is a tridiagonal matrix, explicit formulas for its inverse are available. Applying these formulas gives the expression in the text. □

C CONTINUOUS TALENT-CONTINUOUS TASK MODEL

In this appendix, we briefly describe the continuous talent (and continuous task) assignment model and its optimal control formulation. In our quantitative work, we treat the data as a discrete approximation to this model and solve it using the open-source numerical optimal control software GPOPS-II. Workers are now distributed across an interval of talents $k \in [\underline{k}, \overline{k}]$ according to a distribution function $\Pi : [\underline{k}, \overline{k}] \to [0,1]$ with strictly positive and continuously differentiable density π. As before there is a continuum of tasks ranked by complexity $v \in [\underline{v}, \overline{v}]$. The productivity of talent-task combinations is given by a function $a : [\underline{k}, \overline{k}] \times [\underline{v}, \overline{v}] \to \mathbb{R}_{++}$ satisfying the following assumption.

Assumption 2. *(i) a is twice continuously differentiable on the interior of $[\underline{k}, \overline{k}] \times [\underline{v}, \overline{v}]$ with first derivatives a_i, $i \in \{k, v\}$ and second derivatives a_{ij}, $i, j \in \{k, v\}$. (ii) (strict absolute advantage) $a_k > 0$, (iii) (strict comparative advantage, log supermodularity) $\frac{\partial^2 \log a}{\partial k \partial v} > 0$.*

Otherwise technologies and preferences are as in the main text. An allocation is a triple of measurable functions $c : [\underline{k}, \overline{k}] \to \mathbb{R}_+$, $\nu : [\underline{k}, \overline{k}] \to [\underline{v}, \overline{v}]$ and $e : [\underline{k}, \overline{k}] \to \mathbb{R}_+$ describing the consumption, task and effort assignments of each talent type.[49] As before, task output is linear in labor input. The task output density $y : [\underline{v}, \overline{v}] \to \mathbb{R}_+$ satisfies for

[49]The implicit assumption that all talents are assigned to a specific consumption, task and effort is without loss of generality. It may be shown, along the lines of Proposition 2, that assignment of talents to tasks is strictly increasing in talent given strict comparative advantage.

all k,

$$\int_{\underline{v}}^{\nu(k)} y(v)dv = \int_{\underline{k}}^{k} a[k', \nu(k')]e(k')\pi(k')dk'. \tag{C.1}$$

If ν is differentiable with derivative ν_k, then (C.1) can be re-expressed as, for all k:

$$y(\nu(k)) = a[k, \nu(k)]e(k)\frac{\pi(k)}{\nu_k(k)}, \tag{C.2}$$

Heuristically, the numerator is total output of type k, while the denominator gives the tasks over which the type k workers are "spread". The shadow wage is given by:

$$w[k, v] = b(v)\left(\frac{y(v)}{Y}\right)^{-\frac{1}{\varepsilon}} a[k, v].$$

We restrict planners and policymakers to smooth allocations and mechanisms. This permits the application of optimal control techniques.

Optimal Control Formulation of Government's Problem We formulate the government's problem as a mechanism design problem and recover optimal taxes from this. Mechanisms are analogous to those considered previously in Appendix A. Each worker reports its talent k and, conditional on this, is assigned a consumption c, task ν and effort e. The combination of mechanism and truthfully reported talent imply utility and normalized shadow wage and income levels for each type:

$$\begin{aligned}
\psi(k) &= U(c(k), e(k)) \\
\varphi(k) &= w[k, \nu(k)]/Y^{\frac{1}{\varepsilon}} \\
\rho(k) &= \varphi(k)e(k).
\end{aligned}$$

In addition, let $\omega(k, v) = w[k, v]/Y^{\varepsilon}$. A worker claiming to be type k' must reproduce the observable income level $\rho(k')$. Incentive-compatibility thus requires for all k, k' and v':

$$U(c(k), e(k)) \geq U\left(c(k'), \frac{\rho(k')}{\omega(k, v')}\right). \tag{C.3}$$

Let $\mathcal{U} = \{(u, e) \in \mathbb{R} \times [0, \bar{e}] : u = U(c, e) \text{ for some } c \in \mathbb{R}_+\}$ and let $C : \mathcal{U} \to \mathbb{R}_+$ be defined according to $u = U(C[u, e], e)$. The next proposition gives simpler necessary and sufficient conditions for incentive-compatibility.

Proposition C.1. *Let (ν, e, c) be a smooth mechanism that induces a smooth task output function y. The mechanism is incentive-compatible if and only if: (i) (Monotonicity) $v_k \geq 0$ and $\rho_k \geq 0$*

hold and (ii) (Envelope) the envelope conditions for utility and shadow wages hold:

$$\psi_k(k) = -U_e\left(C[\psi(k), e(k)], e(k)\right) e(k) \frac{a_k[k, \nu(k)]}{a[k, \nu(k)]} \tag{C.4}$$

$$\varphi_k(k) = \varphi(k) \frac{a_k[k, \nu(k)]}{a[k, \nu(k)]}. \tag{C.5}$$

Proof. (*Necessity*). Let (ν, e, c) be a smooth incentive-compatible mechanism. Incentive compatibility implies that $w[k, \nu(k)] \geq w[k, \nu(k')]$ and $w[k', \nu(k')] \geq w[k', \nu(k)]$. Since $w[k, v] = b(v)\ (y(v)/Y)^{-\frac{1}{\varepsilon}}\ a[k, v]$ and a is strictly log supermodular, it follows that if $k > k'$, then $\nu(k') > \nu(k)$. Hence, ν is increasing. To verify (C.5), we apply (envelope) Theorem 4.3 of Bonnans and Shapiro (2000) to: $\max_{[\underline{v}, \overline{v}]} U\left(c(k'), \rho(k')/\omega(k, v')\right)$. This requires U to be continuously differentiable, $\omega(\cdot, v)$ to be continuously differentiable and $\omega(k, \cdot)$ to be continuous. The first two properties hold by assumption (and the definition of ω and w, see (C.2)), the latter holds if y is continuous. Suppose that y is discontinuous at v and, without loss of generality assume $y(v) > y(v_n)$ for some sequence $v_n \to v$. Let $k_n = \nu^{-1}(v_n)$, then for n large enough, $w[k_n, v_n] < w[k_n, v]$, which is a contradiction. Then Theorem 4.3 and Remark 4.14, p.273-4 in Bonnans and Shapiro (2000) and the definition of w imply that the function φ, $\varphi(k) = \max_{v \in [\underline{v}, \overline{v}]} w[k, v]$, is differentiable with $\varphi_k = \varphi \frac{a_k}{a} > 0$.

Let $\rho(k) = \varphi(k) e(k)$ and:

$$\Upsilon[k, k'] = U\left(c(k'), \frac{\rho(k')}{\varphi(k)}\right).$$

Incentive-compatibility requires that: $\Upsilon[k, k] \geq \Upsilon[k, k']$ and $\Upsilon[k', k'] \geq \Upsilon[k', k]$. Hence, $U\left(c(k), \frac{\rho(k)}{\varphi(k)}\right) - U\left(c(k'), \frac{\rho(k')}{\varphi(k)}\right) \geq U\left(c(k), \frac{\rho(k)}{\varphi(k')}\right) - U\left(c(k'), \frac{\rho(k')}{\varphi(k')}\right)$. The assumed Spence-Mirrlees condition and the increasingness of φ, then imply that ρ and c are increasing also. Additionally, since (ν, e, c) is continuous by assumption and w is continuous, Theorem 4.3 in Bonnans and Shapiro (2000) can again by applied to show that: $\psi(k) = \max_{k' \in [\underline{k}, \bar{k}]} \Upsilon(k, k')$ is differentiable with:

$$\psi_k(k) = -U_e\left(C[\psi(k), e(k)], e(k)\right) e(k) \frac{\varphi_k(k)}{\varphi(k)} = -U_e\left(C[\psi(k), e(k)], e(k)\right) e(k) \frac{a_k[k, v(k)]}{a[k, v(k)]}.$$

Sufficiency. Let (ν, e, c) be a smooth mechanism satisfying the conditions in the proposition. The definition of φ, the envelope condition for wages (C.5) and the smoothness of ν imply the first order condition: $w_v[k, \nu(k)]\nu_k = 0$. The smoothness of the various functions also implies that w_v exists and is given by:

$$w_v[k, v] = \left\{\frac{b_v(v)}{b(v)} - \frac{1}{\varepsilon}\frac{y_v(v)}{y(v)} + \frac{a_v[k, v]}{a[k, v']}\right\} w[k, v] \tag{C.6}$$

An worker's optimization over v and k' is separable: regardless of the report choice of k', it is optimal for the worker to select a task v that maximizes its wage $w[k, v]$. Let k^* denote a non-decreasing measurable selection from v^{-1}. Then, using (C.6), the first order condition $w_v[k,\nu(k)]\nu_k = 0$ and log supermodularity, for $\hat{v} > v(k)$,

$$
\begin{aligned}
w[k,v] - w[k,\nu(k)] &= \int_{\nu(k)}^{v} w_v[k,v']dv' \\
&= \int_{\nu(k)}^{\hat{v}} \left\{ \frac{b_v(v')}{b(v')} - \frac{1}{\varepsilon}\frac{y_v(v')}{y(v')} + \frac{a_v[k,v']}{a[k,v']} \right\} w[k,v']dv' \\
&= \int_{\nu(k)}^{\hat{v}} \left\{ -\frac{a_v[k^*(v'),v']}{a[k^*(v'),v']} + \frac{a_v[k,v']}{a[k,v']} \right\} w[k,v']dv' < 0.
\end{aligned}
$$

and similarly for $\hat{v} < v(k)$. Consequently, the mechanism induces a k-worker to choose the task assignment $v(k)$.

Let $k_2 > k_1$, then by the envelope condition for wages, for $k' \in [k_1,k_2]$, $\varphi(k') = w[k',\nu(k')] \geq w[k_1,\nu(k_1)] = \varphi(k_1)$. Combined with the monotonicity and concavity of U, this implies $-U_e\left(c(k'),e(k')\right)e(k') + U_e\left(c(k'),\frac{\varphi(k')}{\varphi(k_1)}e(k')\right)\frac{\varphi(k')}{\varphi(k_1)}e(k') < 0$. The envelope condition for reports and the smoothness of the mechanisms imply:

$$
\Upsilon_{\hat{k}}[k,k]\hat{k}_k = \left\{ U_c(c(k),e(k))c_k(k) + U_e(c(k),e(k))e(k)\frac{\rho_k(k)}{\rho(k)} \right\}\hat{k}_k = 0.
$$

The definitions of Υ and ρ and the preceding discussion then imply:

$$
\begin{aligned}
&\Upsilon[k_1,k_2] - \Upsilon[k_1,k_1] = \int_{k_1}^{k_2} \Upsilon_{\hat{k}}[k_1,k']dk' \\
&= \int_{k_1}^{k_2} \left\{ U_c(c(k'),e(k'))c_k(k') + U_e\left(c(k'),\frac{\rho(k')}{\varphi(k_1)}\right)\frac{\rho_k(k')}{\varphi(k_1)} \right\}dk' \\
&= \int_{k_1}^{k_2} \left\{ -U_e\left(c(k'),e(k')\right)e(k') + U_e\left(c(k'),\frac{\rho(k')}{\varphi(k_1)}\right)\frac{w[k',\nu(k')]}{w[k_1,\nu(k_1)]}e(k') \right\}\frac{\rho_k(k)}{\rho(k)}dk' \leq 0.
\end{aligned}
$$

A similar inequality obtains for $k_1 > k_2$ and so the mechanism induces a k-worker to make a truthful report k. □

It is convenient to define:

$$
\xi(k) := \int_{\underline{k}}^{k} \left(\frac{\pi(k')a[k',\nu(k')]e(k')}{\nu_k(k')} \right)^{\frac{\varepsilon-1}{\varepsilon}} b(\nu(k'))\nu_k(k')dk' \tag{C.7}
$$

and

$$
\zeta(k) = \int_{\underline{k}}^{k} C[\psi(k'),e(k')]\pi(k')dk'. \tag{C.8}
$$

Together ψ, φ, ξ and ζ along with ν form a set of state variables for the optimal control formulation of the planning problem with private information. The envelope conditions (C.4) and (C.5) supply laws of motion for ψ and φ. Equations (C.9) and (C.10) give laws of motion for ξ and ζ:

$$\xi_k(k) = e(k)\varphi(k)\pi(k) \tag{C.9}$$

$$\zeta_k(k) = C[\psi(k), e(k)]\pi(k). \tag{C.10}$$

Finally, the definition of $\varphi(k)$ implies a law of motion for ν:

$$\nu_k(k) = \left(\frac{\varphi(k)}{b(\nu(k))a[k,\nu(k)]}\right)^{\varepsilon} \pi(k)a[k,\nu(k)]e(k), \tag{C.11}$$

The monotonicity conditions on mechanisms needed to ensure incentive-compatibility are omitted and checked ex post. The effort function e is the control. The government's problem becomes:

$$\max_{\psi,\varphi,\zeta,\xi,\nu,e} \int_{\underline{k}}^{\overline{k}} \psi(k)g(k)dk \tag{C.12}$$

subject to the laws of motion (C.4), (C.5) and (C.9) to (C.11) and the boundary constraints:

$$\zeta(\overline{k}) \leq \xi(\overline{k})^{\frac{\varepsilon}{\varepsilon-1}}$$
$$0 = \xi(\underline{k}) \quad 0 = \zeta(\underline{k}) \quad \underline{v} = v(\underline{k}) \quad v(\overline{k}) = \bar{v}.$$

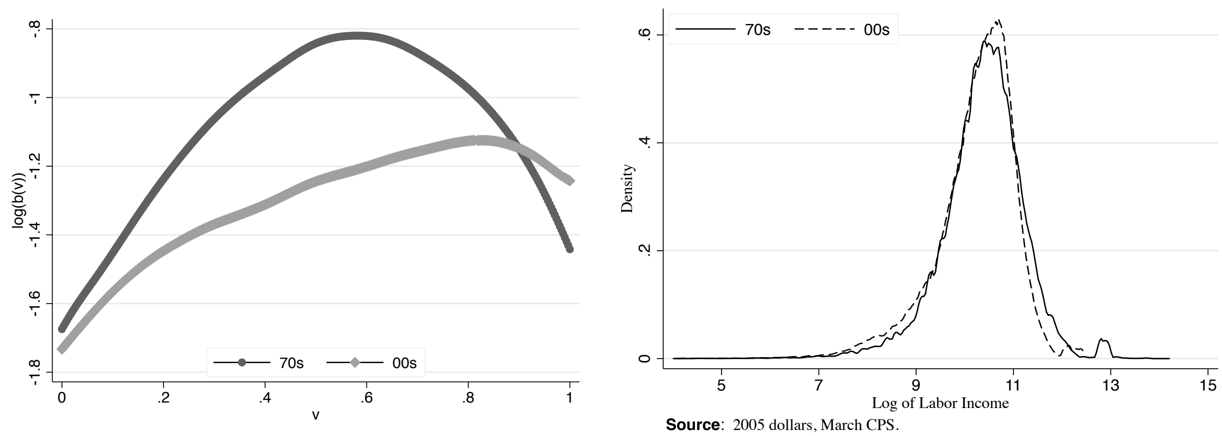
In this problem there is one control (e) and five states $(\psi, \varphi, \zeta, \xi, \nu)$. Routine manipulation of the first order and co-state equations yields the following expression for the optimal effort-consumption wedge:

$$\begin{aligned}
-w[k,\nu^*(k)]\frac{U_c^*(k)}{U_e^*(k)} - 1 = \\
&\underbrace{\mathcal{H}^*(k)\frac{1-\Pi(k)}{\pi(k)}\frac{\varphi_k^*(k)}{\varphi^*(k)}\int_k^\infty \left(1 - \frac{g(t)U_c^*(t)}{p^{\zeta*}\pi(t)}\right)\frac{U_c^*(k)}{U_c^*(t)}\mathcal{N}^*(k,t)\frac{\pi(t)}{1-\Pi(k)}dt}_{\text{Mirrlees}} \\
&\underbrace{-\mathcal{I}^*(k)\left[\frac{p_k^{\varphi*}(k)}{p^{\varphi*}(k)} + \frac{a_k[k,\nu^*(k)]}{a[k,\nu^*(k)]}\right]\frac{p^{\varphi*}(k)}{p^{\zeta*}\pi(k)}\left(-\frac{U_c^*(k)}{U_e^*(k)}\right)}_{\text{Wage Compression}},
\end{aligned} \tag{C.13}$$

where $U_x^*(k) := U_x(c^*(k), e^*(k))$ $x \in \{c, e\}$ and similarly for $U_{xz}^*(k)$, $\mathcal{I}^*(k) := -\frac{1}{\varepsilon}\frac{w[k,\nu^*(k)]}{e^*(k)}$ is the elasticity of the k-talent wage with respect to effort holding the task allocation fixed, $\mathcal{H}^* := \{-\frac{U_{ec}^*}{U_e^*} + \frac{U_{ee}^*}{U_e^*}\}e^* + 1$ is $(1+\mathscr{E}_u)/\mathscr{E}_c$, where $\mathscr{E}_u$ and $\mathscr{E}_c$ are, respectively, the uncompensated and compensated labor supply elasticities, $\mathcal{N}^*(k,t) = \exp\left\{-\int_k^t \frac{e^*(s)U_{ce}^*(s)}{\varphi*(s)U_c^*(s)}\right\}$, $p^{\zeta*} = E\left[\int_{\underline{k}}^{\overline{k}} \frac{1}{U_c^*(t'')}\pi(t'')dt''\right]^{-1}$ is the optimal shadow resource multiplier and $p^{\varphi*}$ is the optimal co-state on the shadow wage φ. The Mirrlees and wage compression components are labeled.

D Dataset

Our main data source is the Current Population Survey (CPS) administered by the US Census Bureau and the US Bureau of Labor Statistics. We focus on the March release of the survey.[50] Data is available continuously from 1968 to 2012. On average each year of data contains about 150,000 observations, from 2001 the sample size has increased to approximately 200,000. The CPS contains detailed information on the demographic and work characteristics of each individual. For additional details on the CPS refer to Heathcote, Perri, and Violante (2010) and Acemoğlu and Autor (2011). The CPS data includes a self reported estimate of hours worked from 1976 onwards. This question as well as questions on income are for the previous calendar year. Hence our sample covers the years 1975 to 2011 (interviews from 1976 to 2012). In the body of the paper we group observations in two groups. We call "the 70s" observations relating to years 1975-1979 (i.e interviewed in years 1976-1980), we call the "00s" observation relating to years 2000-2011 (interviews in 2001-2012).



(a) Values of $\log(b(v))$ over time.

(b) Distribution of log-labor income.

Figure D.1: Estimates on entire sample.

The model analyzed is highly stylized. In order to make data and model compatible (and to reduce the likelihood of measurement error) we further restrict our sample. We drop individuals for whom income, age, sex, education, sector, occupation is not reported. We consider individuals of working age, i.e. between the ages of 25 and 65. We drop individuals with no formal education and the unemployed. Following Heathcote, Perri, and Violante (2010), we also drop underemployed individuals: those working less than 250 hours per year or earning less than \$100 per year (dropping an additional 196,684 observations). Our final sample comprises of 2,039,123 individual/year observations. All variables are weighted with the provided weights and dollar denominated variables are

[50] Data is taken from: Miriam King, Steven Ruggles, J. Trent Alexander, Sarah Flood, Katie Genadek, Matthew B. Schroeder, Brandon Trampe, and Rebecca Vick. Integrated Public Use Microdata Series, Current Population Survey: Version 3.0. [Machine-readable database]. Minneapolis: University of Minnesota, 2010.

deflated using CPI to 2005 dollars.[51] In Figure D.1b we display the evolution of the distribution of log labor income between the "70s" and the "00s". The main feature that emerges is the widening of the distribution in the "00s" relative to the "70s".

Alternate Sample Selection

We briefly explore the impact of our sample selection on the estimated values of $b(v)$. In Figure D.1a we consider the CPS sample before applying our sample selection (but after removal of individuals with missing information or an unclassifiable occupation). As can be seen polarization is still apparent. However for low v occupations we observe little change between the two decades. Note that a similar result would appear using the sample selection of Acemoğlu and Autor (2011). This is because the authors only remove individuals who worked less than one week in the previous year or are less than 16 years of age.

E Counterfactuals

In this appendix we separately evaluate the impact of change in the a and b functions on optimal policy. To do so, we first hold the parameters of a fixed at their 1970s values, while allowing those of b to change to their 2000s values; we compute the corresponding optimal tax equilibrium. We then repeat the exercise holding the parameters of b fixed, while allowing those of a to change. We compare the resulting tax equilibria to those in which both functions are at their 1970s or 2000s levels.

Assignment Figure E.1 shows the impact of the empirical a and b changes together and isolation on density of workers across tasks. Changes in b alone lead to quite large changes in the relative "number" of workers performing tasks. In particular, the polarizing adjustments in task demand (growth at the extremes relative to the middle) occurring between the 1970s and the 2000s induce growth in the density of workers at the extremes and, hence, job polarization in the associated optimal tax equilibrium. Changes in a alone have an opposite (if more modest) effect: the number of workers performing mid-level tasks grows relative to the extremes. This reflects productivity growth in low tasks by low talents and in high tasks by high talents inducing reductions in shadow task prices at the top and the bottom and movements of some lower and higher talents into mid-level tasks. However, when changes in the b and a parameters are combined, it is the former that dominates.

[51] CPI for all urban consumers, all goods.

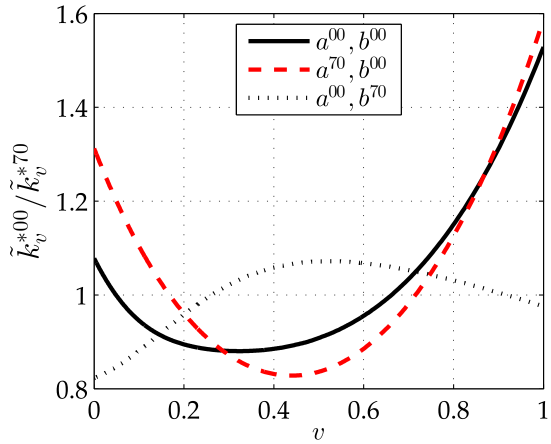


Figure E.1: Relative changes in $\tilde{k}_v^*$ from the 1970s to the 2000s. Allowing b to change, a to change and both a and b to change.

Although, the b parameter change induces quite large changes in the numbers of workers performing particular tasks, this is achieved with only modest occupational re-assignments of given workers. As shown in Figure E.4, low-mid level talents reduce their task assignment, but by no more than 2%, high-mid level talents increase their task assignment, but by no more than 3%.

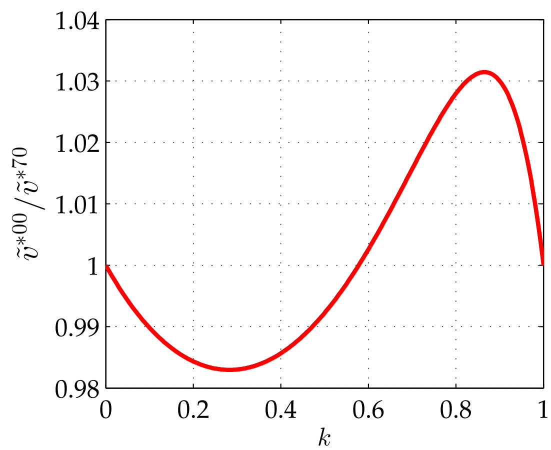


Figure E.2: Relative changes in task assignment $\tilde{v}^*$ induced by the shift in b from 1970 to 2000.

WAGE CHANGES An implication of the modest change in task assignment induced by the shift in the b function is that equilibrium wage growth over talents is also only modestly altered by this shift. As shown in Figure E.3, b changes alone induce very slight compression in wages across low-to-mid talents and very slight stretching and expansion of wages across mid-to-high talents. Changes in the a function also depress wage growth across talents at the bottom and raise it at the top, but the effect is much more pronounced.

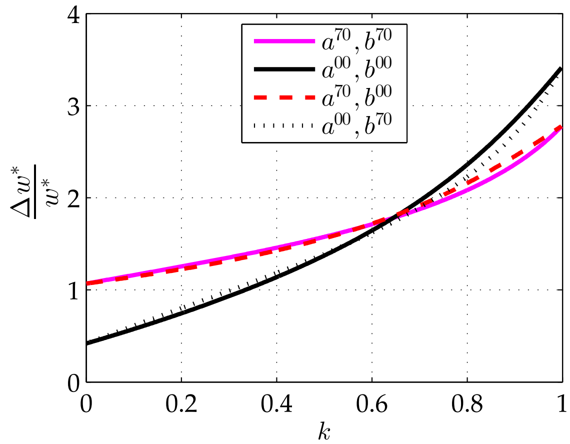
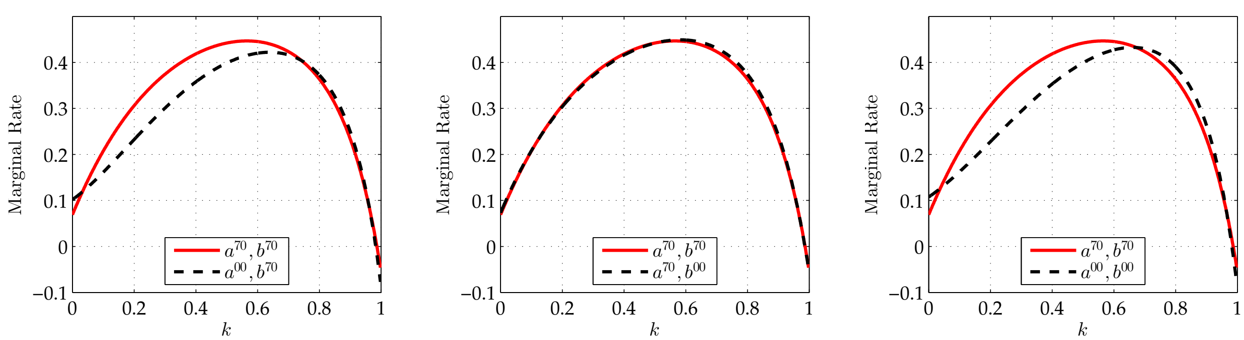


Figure E.3: Equilibrium wage growth $\Delta w^*/w^*$ for the parameter combinations (a_{70}, b_{70}), (a_{70}, b_{00}), (a_{00}, b_{70}) and (a_{00}, b_{00}).

MARGINAL TAX CHANGES The shift in the b function alone has limited impact on the relative wage-effort elasticities and on the wage compression term. Combined with its small impact on wage growth over talents, it has a correspondingly modest effect on marginal taxes, see Figure E.4a. In contrast, the shift in the a function has a much more significant impact on wage growth and on the relative wage-effort elasticities. It has a much more significant effect on optimal marginal taxes and accounts for most of the adjustment between the 1970s and the 2000s.



(a) Effect of changes in b. (b) Effect of changes in a. (c) Effect of changes in a, b.

Figure E.4: Marginal tax effects.