Technological Change and Earnings Inequality in the U.S.: Implications for Optimal Taxation*

Pedro Brinca† João B. Duarte‡ Hans A. Holter§ João G. Oliveira¶

July 7, 2025

Abstract

Since 1980, earnings inequality in the U.S. rose steadily alongside rapid technological change. To what extent does technological change explain the observed increase in earnings dispersion? How does it affect the optimal progressivity of labor earnings? To answer these questions, we develop an incomplete markets model with occupational choice. We estimate an aggregate production function with capital-occupation complementarity and four occupations that differ with respect to cognitive complexity and routine task intensity. We calibrate our model to resemble the U.S. economy in 1980 and find that technological transformation can account for two thirds of the increase in earnings dispersion between 1980 and 2015. The main driver is the rising relative wage of non-routine cognitive occupations, which benefit the most from complementarity with capital. Although technological growth is associated with higher earnings inequality, it leads to a significant drop in optimal tax progressivity. Lower progressivity (in particular when combined with tech change) leads to more capital accumulation and an inflow of workers into higher-paid occupations. This increases output but also raises the wages of the occupations at the bottom of the wage distribution, dampening the redistributive gains from progressive taxation.

Keywords: Earnings Inequality, Taxation, Technological Change, Automation

JEL Classification: E21, E23, E62, H21, H23, J24, J31, O33, O40

---

*We thank Árpád Ábrahám, Daron Acemoglu, David Autor, Vasco Botelho, Juan Dolado, Loukas Karabarbounis, Nick Kozeniauskas, Musa Orak, Lee Ohanian, Kurt Mitman, Felipe Netto, Ana Pereira, Cezar Santos, Pedro Teles, Gianluca Violante, seminar participants at the Lisbon Macro Group, the EUI Economics Department and participants at the PEJ 2018, the 2019 North-American Winter Meeting of the Econometric Society, the 2019 LubraMacro, and the 2023 PSE Macro Days conference, and two anonymous referees for their helpful comments and suggestions. Pedro Brinca, João Duarte and Hans A. Holter are grateful for financial support from the Portuguese Science and Technology Foundation (FCT), grants number PTDC/EGE-ECO/7620/2020, UID/ECO/00124/2013 and UID/ECO/00145/2013, POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences Data Lab, Project 22209), and POR Norte (Social Sciences Data Lab, Project 22209). Pedro Brinca is grateful for his CEECIND/02747/2018 FCT grant, João Duarte for his CEECIND/03227/2018 FCT grant and Hans A. Holter for his CEECIND/01695/2018 FCT grant. João Oliveira is grateful for his FCT grant SFRH/BD/138631/2018. Hans A. Holter thanks the Research Council of Norway, grants: 283314, 302661 (IMWEL), 267428 (OFS), and 341250 (NFS).

†Nova School of Business and Economics.E-mail: pedro.brinca@novasbe.pt

‡Nova School of Business and Economics. E-mail: joao.duarte@novasbe.pt

§University of Delaware, University of Oslo, and Nova SBE. E-mail: hans_holter@hotmail.com

¶Bank of England. E-mail: joao.oliveira@bankofengland.co.uk

1 Introduction

Earnings inequality in the U.S. rose steadily since 1980 (see Figure 1a, left panel).[1] What accounts for the large increase in inequality, and what are the policy implications? There is a heated debate about these questions among academics, policymakers, and the public press. A common view, and perhaps conventional wisdom, is that one should meet increased inequality with higher and more progressive taxes (Delaney, 2017).

Alongside the increase in inequality, there has also been technological progress. Figure 1b displays a rapid fall in the relative price of equipment investment goods, which can be viewed as reflecting Investment-Specific Technological Change (ISTC) such as cheaper access to computing power and storage (Krusell et al., 2000; Karabarbounis and Neiman, 2014). In this paper, we answer the following questions: (i) to what extent does technological change explain the observed increase in earnings inequality? (ii) how does it affect the optimal progressivity of the tax and transfer system?

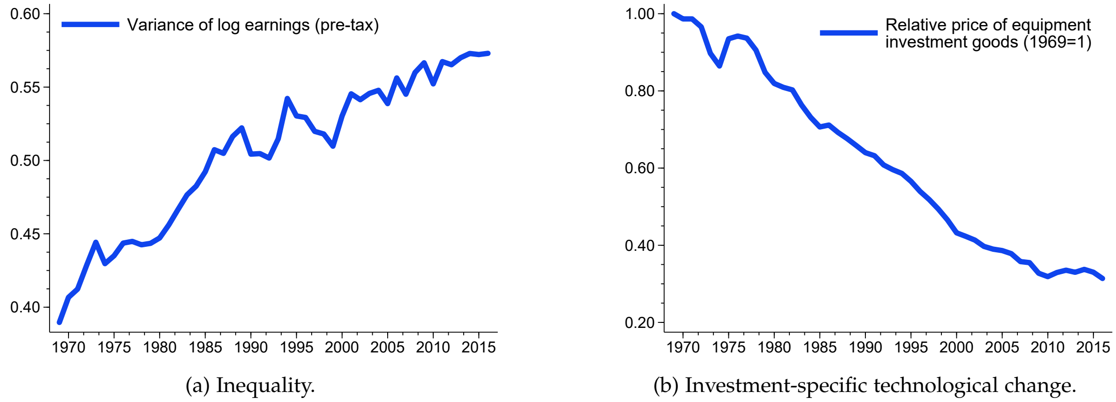
The literature on technological change and the labor market emphasizes task specificity and the degree to which workers' tasks are complementary to capital as crucial determinants of wages. Autor et al. (2003) introduce a framework where occupations differ in terms of the nature of the tasks being performed. There are four main categories of tasks: Non-routine cognitive (NRC), non-routine manual (NRM), routine cognitive (RC) and routine manual (RM). To study the evolution of inequality and the implications for optimal tax policy, we adopt this categorization and develop an incomplete markets, heterogeneous agent model with technological change and occupational choice.

Our first contribution is to expand on the seminal paper by Krusell et al. (2000) by specifying and estimating an aggregate production function with labor inputs based on occupation categories rather than the education levels of the workforce. We provide novel estimates for the elasticities of substitution between structures, equipment capital, and these four occupation categories, which have been extensively used in the literature that studies the impact of technological change on labor markets. This production function is the centerpiece of our model and we study the effects of technological change by inserting time-varying estimates of its parameters.

Second, we are the first to explain the increase in earnings inequality in the U.S. in a framework with technological change by estimating a production function with capital-occupation complementarity. The previous literature focused on the education skill premium or the labor share of earnings, using representative agent frameworks.[2] Changing only the time-varying

[1]Figure 1a shows this phenomenon since 1970. However, our analysis focuses on the period from 1980 to 2015 due to the limited availability of the necessary data before 1980.

[2]Krusell et al. (2000) show how capital-skill complementarity can explain the evolution of the college wage premium, but they do not study other measures of inequality. Eden and Gaggl (2018) focus on the evolution of the labor share and the routine versus the non-routine labor share. Finally, Vom Lehn (2020) studies the relative wages of three types of workers as opposed to our four. Using a nested CES production function, where capital equipment is directly substitutable with routine workers only, and a different calibration strategy, he concludes that technological change cannot account for the labor market polarization in wages.



Note: The variance of log pre-tax earnings is computed from the CPS for employed workers. A description of the data is provided in Appendix A. ISTC is measured by the change in the relative price of equipment investment, computed as the ratio between equipment investment prices from the BEA and the BLS urban consumer price index.

Figure 1: Inequality and ISTC.

estimates of the production function parameters in our model, with endogenous occupation choice, savings and labor supply decisions, we can explain two thirds of the increase in earnings inequality (measured as the variance of log earnings), most of the changes in occupational shares, and most of the changes in the wage premia between our four occupation groups, between 1980 and 2015.

Our third contribution is to investigate the quantitative implications of technological change for optimal labor earnings tax progressivity. Typically, the optimal progressivity of the tax system can be analyzed as a trade-off between redistribution and insurance on the one hand and efficiency on the other. Introducing margins such as human capital or capital production externalities tilts this trade-off in favor of efficiency. But how strong are these effects? We take a more quantitative approach to this question than the previous literature. We use our estimates of production function parameters to calibrate a quantitative incomplete markets model, which has both human capital (through occupational choice) and a production externality from physical capital (through capital-occupation complementarity), and which furthermore succeeds at explaining a large share of the increase in U.S. earnings inequality over time. We show that the technological transformation between 1980 and 2015, particularly ISTC, calls for a significant drop in tax progressivity.

Our framework is, in some respects, a standard life-cycle model with incomplete markets and idiosyncratic risk. On the household side, it is, however, distinguished by a once and forever choice between our four occupations at the beginning of work-life.[3]. Agents select an occupation based on an idiosyncratic cost of acquiring the necessary skills and on the expected lifetime utility from consumption and work effort in each profession. A key element of this choice is the set of occupation wage rates. These are determined at the aggregate level by market forces which respond to technological change, but also at the individual level by a worker's productivity endowment. In particular, households select their occupation

[3]Some workers do, of course, retrain. However, Cortes et al. (2020) provide evidence of the fall in routine employment in the U.S. being primarily caused by declining inflow rates among younger workers.

conditional on their level of permanent ability, creating a role for self-selection.

A major departure from previous literature and of crucial importance for the quantitative results of the paper lies in the aggregate production function we use. Our production function has six inputs: Four occupations, equipment capital and structure capital. Furthermore, there are three sources of technological change: ISTC, latent occupation-biased technological change (LAT) and TFP growth. To quantify the labor inputs in each occupation, we apply the cross-walk classification table developed by Cortes et al. (2020) to map tasks into occupation codes to create a set of four major categories. The extent to which labor demand and wages in each of these occupation categories will affect the wage distribution is determined by their respective roles in the production function, by latent occupation-biased technological change, and, in particular, by their complementarity with equipment capital. For example, the effect of a fall in the price of equipment investment goods (ISTC) is to spur capital accumulation and create increased demand for workers in occupations with tasks that are more complementary to equipment capital relative to those that are less so. Due to mobility frictions and heterogeneous entry costs across occupations, rising labor demand in some occupations generates wage premia relative to others.

We parameterize the model in two steps. First, we use the firms' first-order conditions and a no-arbitrage condition, which restricts the expected net return of equipment to be the same as that of structures, to estimate the production function given data hours worked and returns on structures and equipment capital. We use a simulated pseudo maximum likelihood (SPML) approach, as proposed in Ohanian et al. (1997) and Krusell et al. (2000), that implicitly targets the wage bill ratios of the NRC, NRM and NRM occupations to the RM occupation from 1968 to 2015. We find that our estimates deliver a very good fit to the targeted moments. Second, we insert the estimated production function into our incomplete markets model, and we calibrate the remaining parameters to resemble the U.S. economy in 1980.

To analyze the relationship between technological growth and earnings inequality we study our model during a 100-year transition, starting in 1980. In 1980, the agents suddenly learn the paths of changing prices and taxes over the next 100 years (i.e. they receive an MIT shock). We let a number of parameters change exogenously between 1980 and 2015 in our baseline experiment. This includes our estimated technology parameters, social security, taxes on consumption, capital, and labor, and government debt. We assume that the parameters of the model change until 2015 and are fixed after that, i.e, we make no projections about future technological change or taxes.

We find that the technology parameters are the relevant ones for explaining the increase in earnings inequality, and the changes in occupational shares and wage premia. Inserting only our time varying estimates of the production function parameters governing (i.e., ISTC, LAT, and TFP) we find that technological change (in particular LAT and ISTC) can account for about two thirds of the increase in earnings inequality and most of the changes in occupation shares and wage premia between 1980 and 2015.[4] The ability of the model to explain the

[4]This finding is consistent with Barro (2000) who finds that across rich counties, inequality and economic

evolution of these statistics over time depends on the endogenous occupation choices, savings and labor supply responses of the agents in our model after treating it with the changes in the price of equipment (ISTC) and the levels of LAT and TFP between 1980 and 2015. The main driver of the increase in earnings inequality is the rising relative wage of non-routine cognitive occupations, which benefit the most from complementarity with capital. Changes in LAT are the single largest contributor to the rise in the variance of log earnings, accounting for 41% of the increase. Combined with ISCT, they explain 65% of the growth in earnings inequality between 1980 and 2015.

We validate the model's explanation of rising inequality through a purely empirical exercise (see Section 6.2). We show that changes in wage premia and shifts in occupation employment shares account for 68% of the increase in earnings inequality between 1980 and 2015 — two forces that also change endogenously in the model as a result of technological progress and household choices, yielding a nearly identical figure.

Since the model accounts for a large share of the rise in earnings inequality and replicates the observed responsiveness of occupational choices to technological change, we consider it well suited for quantitative analysis of optimal tax policy. We further validate its suitability by confirming that the elasticity of occupational choice to changes in tax progressivity aligns with empirical estimates (see Appendix G.1).

We study optimal taxation both in a long-run steady state and taking into account a transition starting in 1980.[5] Our optimal steady-state tax experiment is to maximize the expected welfare of an unborn individual with respect to the progressivity and level of the labor income tax code, taking government expenditure and other taxes as exogenously given.[6] When taking the transition into account, we redefine the welfare criterion as the sum of the expected discounted utility of each generation entering the labor market in every period, starting from 1980. We then study the interaction between optimal tax progressivity and our three sources of technological change, and use the framework of Flodén (2001) (see also Benabou, 2002 for a similar approach) to decompose the welfare effects of progressive taxation into the contributions resulting from its impact on efficiency, redistribution and insurance.

We apply a non-linear tax function as in Benabou (2002) and Heathcote et al. (2017), $y_a = 1 - \theta_0 y^{-\theta_1}$, where y_a denotes after-tax income and θ_0 and θ_1 define the level and progressivity of the tax system, respectively. For 1980, we find the optimal value of our measure of tax progressivity, θ_1, to be 0.15 in a long-run steady state (below the estimated benchmark value of 0.19). Replacing the 1980 technology parameters with their 2015 counterparts, we find that a value of 0.03 is optimal.[7] To give an interpretation in terms of actual tax rates: The

growth are correlated.

[5] We do this both for comparison with the literature and to be transparent about short- and long-run considerations.

[6] This is the classic tax experiment in the literature on incomplete market models with heterogeneous agents.

[7] Indeed, there is evidence of some reduction in tax progressivity in the U.S. since 1980. Wu (2021) finds that this measure of progressivity has fallen from 0.19 to 0.14 between 1980 and 2015. In Section 7.3 we show that whereas technological change calls for much flatter taxes some of the other factors that changed between 1980 and 2015 call may have offset this effect.

average tax rate for an individual with Average Earnings (AE) is 15% both with $\theta_1 = 0.15$ and $\theta_1 = 0.03$. The average tax rates for two individuals making 0.5AE and 2AE are, however, 5.7% and 23.4% with $\theta_1 = 0.15$ and 13.2% and 16.7% with $\theta_1 = 0.03$.

The main mechanisms driving this result are the rising productivity of NRC professions, the positive effect of shifting workers to NRC occupations on the wages of lower-paid occupations, and the higher returns to wealth with the 2015-technology.[8] Reducing tax progressivity shifts workers towards higher-paying occupations, which raises output as well as the wages in lower-paying occupations, but also reduces the benefits of redistribution and insurance from the tax system.[9] This tradeoff is, however, tilted towards flatter taxes when we account for the technological transformation between 1980 and 2015.

Accounting for the transition generally leads to higher optimal tax progressivity. The positive effects from technological change and lower tax progressivity take time to materialize and the first generations in the transition, tilting the trade-off in favor of more progressive taxes. Making a once and for all change in tax progressivity in 1980, while starting the technological transition and letting the tax level adjust to clear the budget in every period, we find an optimal progressivity value of $\theta_1 = 0.1$. Finally, letting optimal progressivity vary over time between 1980 and 2015, we find a high starting progressivity value of $\theta_1^{1980} = 0.42$ and a low end value of $\theta_1^{2015} = 0.05$.[10]

Among our three sources of technological change, ISTC is the main force responsible for the drop in optimal tax progressivity (LAT and TFP works to complement the effect of ISTC). From the perspective of the social planner, all three welfare impacts of progressive taxation (efficiency, redistribution and insurance) are tilted towards lower optimal progressivity with higher ISTC. First, the efficiency channel is stronger because there is more capital and stronger complementarity with high-earning occupations. The benefit from lowering the marginal tax rates on high earners and getting people to select NRC professions is thus higher. Second, although there is more earnings inequality in 2015, which creates additional incentives for redistribution, more agents moving from low-earning to high-earning occupations increases the wage rates of low earners and decreases the wage rates of high earners. The positive effects that people moving to high-earning occupations have on the wages of low-earning occupations dampens the redistribution channel loss from flatter taxes. Finally, ISTC is responsible for the increased returns on capital in 2015, which dampens the insurance motive. A higher return on capital makes it cheaper to self-insure and weakens the insurance role of a progressive tax system.

The rest of the paper is organized as follows. Section 2 contains a brief survey of the related literature. In Section 3, we describe the model. In Section 4, we estimate the aggregate production function. Section 5 is devoted to calibrating our model. In Section 6 we present

[8] See Jordà et al. (2019) for evidence of higher return rates on wealth in the U.S. Moll et al. (2022) also argue that technological change raises the return on wealth.

[9] Without occupational choice there is only a slight drop in optimal progressivity between 1980 and 2015.

[10] Guerreiro et al. (2021) find a similar result for time varying optimal capital taxation in a model with automation. It is optimal to tax robots in the short run but not in the long run.

the results on technological transformation and the change in earnings inequality, and in 7 we study optimal taxation. We conclude in Section 8.

2 Relation to the Literature

This paper relates to two main strands of literature. First, the literature investigating the impact of technological change on wages and inequality. Second, the literature on optimal Ramsey taxation in incomplete markets models with heterogeneous agents.

Our work builds on the classic paper by Krusell et al. (2000). We expand their framework by specifying and estimating an aggregate production function with labor inputs based on occupations rather than the education levels of the workforce. Krusell et al. (2000) document the impact of skill-biased technological change and capital-skill complementarity on the skill premium (i.e., the college premium) and can explain its evolution over time using this mechanism. Their approach is, however, a purely production-side approach with two types of labor (high-skilled and low-skilled). They do not model households' endogenous savings and labor supply decisions, and they do not study broader inequality measures, such as the variance of log earnings.[11] Using our framework with four types of labor but also rich agent heterogeneity in the forms of income risk, age, savings and permanent ability, we can explain the changes in wage premia between our four occupation groups, two thirds of the increase in earnings inequality in the U.S., measured as the variance of log earnings, as well as the change of occupation shares between 1980 and 2015. This result depends on our estimation of the production function and on the endogenous occupation choices, savings, and labor supply decisions of the agents in our model in response to the changes in the price of equipment (ISTC) and the levels of LAT and TFP.

Instead of dividing the population by education level, Autor et al. (2003) argues that the most empirically relevant interaction between technology and worker productivity comes from the types of tasks a worker performs (although these are correlated with education). They study the effect of computerization on changes in employment by occupation categories and posit that some occupations have a prevalence of tasks that can easily be automated and solved by machines (routine tasks). In contrast, others involve complex problem-solving and interactions (so-called non-routine tasks) which are very costly or impossible to automate. The other key distinction of tasks is whether they are cognitive or manual. We adopt the occupation taxonomy of Autor et al. (2003) and use the cross-walk classification table developed by Cortes et al. (2020) to map tasks into occupation codes to calculate equilibrium quantities of labor input by occupation category.[12]

There is a growing literature classifying labor inputs by tasks and studying the interaction

[11] Slavik and Yazici (2022) make the point out that an increase in idiosyncratic earnings risk over time has contributed to the increase in the skill premium by inducing precautionary savings behavior, which through capital-skill complementarity drives up the skill premium.

[12] See Appendix A for additional details on data treatment. We use these data to construct time series on employment and wages by occupation category. To calculate wage premia, we use the method of Krusell et al. (2000), as described in Appendix B.

with automation technologies. Eden and Gaggl (2018) also estimate an aggregate production function for the U.S. using the routine/non-routine paradigm and investigate the welfare implications of investment-specific technological change for the welfare of a representative agent. Our work instead uses the four task dimensions postulated by Autor et al. (2003). Also, it allows for labor-augmenting technological change at the occupation level, which will be important for our findings below showing that workers at the bottom of the wage distribution have enjoyed wage growth relative to the center of the distribution as a result of technological change. Vom Lehn (2020) maps tasks into three labor types and proposes an aggregate production function that is closer, but still quite different from our production function specification.[13] In contrast to our results, he finds that his calibrated model cannot reproduce the job market polarization in wages. The difference between his findings and ours possibly stems from the different production function specifications, the calibration procedures or the different classifications of labor inputs. Other papers using a task-based framework to study the impact of technological change on inequality include Acemoglu and Autor (2011), Acemoglu and Restrepo (2018), Kaplan and Zoch (2020), and Moll et al. (2022). We do not follow some of these studies in modeling tasks explicitly. We thus forego a more detailed characterization of the production process in favor of the ability to measure the inputs in production more accurately, enabling the estimation of the production technology in Section 4 below.

This paper is also related to the literature on optimal progressive Ramsey taxation in incomplete markets models with heterogeneous agents. Typically, it has focused on maximizing welfare in long-run steady states, including Conesa et al. (2009), Peterman (2016), Heathcote et al. (2017), Heathcote et al. (2020), and Wu (2021). However, recently it starting considering transitions after once and forever tax changes (e.g., Bakis et al., 2015, Kindermann and Krueger, 2022, Boar and Midrigan, 2022, Ferrière et al., 2023, Kina et al., 2024), and optimal dynamic taxation during a transition (e.g., Dyrda and Pedroni, 2021, Acikgoz et al., 2022). For comparison with these different strands, and to understand the impact of short-run versus long-run effects of technological change on optimal taxation, in Section 7 we analyze optimal tax progressivity: (i) in the long-run; (ii) in a transition after a once and forever tax change; and (iii) allowing for time-varying tax progressivity (see Section 7.3).[14] All of the above papers with transitions (except for Kina et al., 2024) do, however, have in common that they work with the classical Aiyagari (1994) model. Our contribution is to quantify the impact of technological change and human capital (through occupational choice) on optimal tax progressivity. These are two factors of crucial importance to inequality as well as the trade-offs between efficiency, redistribution, and insurance that one must consider when designing optimal tax systems.

A subset of these studies focused on the question of how the tax system should respond to

[13]In his production function, abstract and manual labor inputs are substitutes or complement to a bundle composed of routine labor input and capital equipment. In contrast, in our framework, NRC, NRM and RC all have a constant elasticity of substitution with capital equipment directly.

[14]The approach of Acikgoz et al. (2022) allows for a different tax rate in every period of the transition. Due to our much richer model, we take a simpler route and let the tax rate transition linearly between 1980 and 2015.

increasing inequality caused by various sources. Closest to ours are Wu (2021) and Heathcote et al. (2020). Wu (2021) considers an aging population, shrinking gender wage gap, increased idiosyncratic risk, and increased skill premium (modeled with a parameter governing the returns to human capital investment). In total, these changes lead to a slight drop in optimal tax progressivity. The effect of an increase in the skill premium on optimal progressivity is, however, almost neutral. Heathcote et al. (2020) study the impact of technological change on optimal progressivity in an incomplete markets model with a continuous skill choice. They also find that skill-biased technological change has limited, downward, impact on optimal tax progressivity. However, their focus is on college education and skill-biased technological change. There is no role for capital in production and no occupation choice. Our paper takes an occupation-based approach and focuses on the role of capital-occupation complementarity. In contrast to these two studies, we find a striking drop in optimal tax progressivity due to ISTC.[15]

Related to our work is also Ales et al. (2015) who study Mirrlesian taxation in a static, talent assignment model without capital but with technical change. They find that technical change should lead to a slightly more progressive tax system. Kina et al. (2024) study optimal capital taxation in a model with two skill levels and capital-skill complementarity. They find that the presence of capital-skill complementarity leads to more need for redistribution and a higher level of capital tax. Guerreiro et al. (2021), study optimal robot taxation in a model with the possibility of automation of tasks and endogenous choice between two occupations.[16] They find that one should tax robots in the short-run, to reduce inequality, but not in the long-run, when workers can choose occupation freely. Our contribution is distinct from theirs in that we focus on the labor income tax, which affects occupation choice incentives directly, and estimate the production function which governs the technology process. We also broaden the analysis to include the cognitive/manual dimensions of tasks, and we include idiosyncratic income risk (adding an insurance motive to the optimal taxation problem).

3 A Model of Labor Market Inequality and Technological Change

Our model is a life-cycle version of the Bewley-Aiyagari-Hugget model:[17] An incomplete markets economy with overlapping generations of heterogeneous agents and partially uninsurable idiosyncratic risk that generates income and wealth distributions.

Before entering the labor market, households choose an occupation. This decision is the result of a cost-benefit analysis depending on the idiosyncratic cost (or benefit) of acquiring

[15] This large effect is consistent with the argument in Powell and Shan (2012), who show that it is progressivity and not the level of taxation which is relevant for occupation choice. In our case this effect is compounded by the fact that occupational choice is irreversible. We view this modeling choice as more realistic in our case, given the difficulty and cost of retraining workers to perform non-routine cognitive occupations halfway through their lives.

[16] Like us both Guerreiro et al. (2021) and Kina et al. (2024) assume that older generations cannot change occupations This is in line with the evidence provided by Cortes et al. (2020), who argue that the fall in routine employment in the U.S. has been primarily caused by declining inflow rates among younger workers.

[17] See Bewley (2000), Aiyagari (1994), and Hugget (1993).

the necessary skills to perform it and the expected earnings in that occupation conditional on the agent's permanent ability. For tractability, we assume the decision is irreversible and mutually exclusive, and determines in which labor market the household will be working for the duration of its working life.[18] After choosing an occupation, households face a stream of idiosyncratic wage shocks and make joint decisions about consumption, savings and hours worked.

For the production side of the economy, we draw on the modeling strategy of Krusell et al. (2000) and Karabarbounis and Neiman (2014). There are three final goods sectors in the economy: the consumption goods, structure capital goods, and equipment capital goods sectors. This formulation allows us to express the price of equipment goods as a function of the level of technology in that sector relative to the consumer goods sector.

The centerpiece of the model is the production function for the intermediate input, which uses a combination of the different occupation and capital types to produce final goods. We build on the production function introduced by Greenwood et al. (1997) and extend it to encompass four labor varieties: Non-routine cognitive, non-routine manual, routine cognitive, and routine manual.

Technological progress, in the form of total factor productivity growth, occupation-biased technological change, and investment-specific technological change, affects capital and labor demand and occupation wage premia. This framework creates a rich interaction between capital accumulation, technological change, and the wages of different occupations and allows us to map the dynamics of these variables into earnings inequality measures.

In Section 6, we use a calibrated version of the model to measure the fraction of the observed increase in earnings inequality between 1980 and 2015 which is attributable to technology, and in Section 7 we use it as a laboratory to study optimal tax policies and how they depend on technological change. Below, we describe the household problem, the production side of the economy, and the definition of equilibrium in more detail.

3.1 Demographics and Occupational Choice

The economy is populated by $J = 81$ overlapping generations. A period in the model corresponds to one year, and households begin life at age 20. Thus, household age, j, varies between 1 (for age 20 households) and 81 (for age 100 households). Households differ with respect to their occupation, o_i, persistent idiosyncratic productivity shock, u_{ij}, permanent ability, a_i, and asset holdings, b_{ij}. Working age agents choose how much to work, h_{ij}, how much to consume, c_{ij}, and how much to save, b_{ij+1}, to maximize utility.

Before joining the labor market at age $j = 1$, households make an irreversible and mutually exclusive occupation choice. They do so conditional on an idiosyncratic taste shock, which defines the cost (or benefit) of acquiring the skills necessary to each occupation; on their ability,

[18] Cortes et al. (2020) provide evidence of the main driver of the decline in routine employment being a reduction in inflow rates rather than an increase in outflow rates. This is consistent with our assumption of inability to change occupation type during work life, despite changing wage premia in other occupation types.

which may have a different pecuniary return across occupations; and on expected occupation wage rates and occupation-specific distributions of idiosyncratic productivity shocks.

Thus, a household i draws taste shock, κ_{io}, from joining occupation type $o \in O = \{\text{NRC}, \text{NRM}, \text{RC}, \text{RM}\}$. This term can be viewed as the personal cost (or benefit, if positive) of acquiring skills to perform the tasks associated with a given occupation, such as the effort (or joy) from studying in the case of cognitive occupations, for example. We assume κ_{io} follows a type 1 extreme value distribution with location parameter $\mu_{\kappa,o}$ and scale parameter $\sigma_{\kappa,o}$ in the tradition of discrete choice modeling as in McFadden (1973).[19]

Denote $V(1, b_{i1}, o, a_i, u_{i1})$ as the expected discounted lifetime utility of household i in occupation $o \in O$ at labor market entry ($j = 1$), with starting bond endowment b_{i1}, ability a_i, and starting level of idiosyncratic productivity risk u_{i1}. Permanent ability is drawn randomly from a mean-zero standard normal distribution and is known to agents when they choose their occupation. They start their life with no savings ($b_{i1} = 0, \forall i$), and u_{i1} is drawn randomly from the stationary distribution of the Markov process u_i. Denote household i's expected value (taken over u_{1i}) of joining occupation $o \in O$ as $V_o(a_i)$. The occupation choice problem is given by:

$$\max_{o \in O} \quad U_{io} = V_o(a_i) + \kappa_{io}, \tag{1}$$

where $\kappa_{io} = \mu_{\kappa,o} + \sigma_{\kappa,o} v_{io}$ is an occupation-specific idiosyncratic taste shock, and v_{io} is a mean-zero standard Gumbel random variable. As is standard in the literature, we assume independence between taste shocks across occupations and that $\sigma_{\kappa,o} = \sigma_\kappa$, i.e., a common scale parameter between all shocks.[20] Households choose the occupation where total utility is highest given a_i. This specification of the problem leads to self-selection of agents into occupations (in the tradition of Roy, 1951), affecting effective labor supply and within-occupation wage dispersion, due to differences in the return to ability across occupations (see section 3.3).

After retiring at age 65 (model age 46), households face an age-dependent probability of dying, π_j, dying with certainty at age 100. $s_j = 1 - \pi_j$ defines the age-dependent probability of surviving, so that in any given period, using a law of large numbers, the mass of retired agents alive at model age $j > 45$ is equal to $S_j = \prod_{l=46}^{l=j} s_{l-1}$.

Dying households leave bequests which are redistributed evenly in a lump-sum manner between the households that are currently alive, denoted by Γ. We include a bequest motive in this framework to make sure that the age distribution of wealth is empirically plausible, as in Brinca et al. (2021).

Retired households make consumption and saving decisions and receive a retirement benefit, $\Psi_t(o_i, a_i)$. For simplicity, we assume that the public retirement benefit is equal to a fraction,

[19]Concretely, this formulation is the same as that used for unordered multinomial models where discrete choices are modeled as outcomes from an additive random utility model. See Cameron and Trivedi (2005) for a detailed exposition. Additionally, we adopt the convention that names and indices (i.e. occupation, age, time, individual) in subscripts are separated by a comma.

[20]See Boar and Lashkari (2021) and Guerreiro et al. (2021), for example. These assumptions guarantee that employment shares are well behaved and location parameters are unique for a given normalization of the location parameter of the benchmark economy (RM in our case). See McFadden (1973) for a detailed derivation.

$\psi_{ss,t}$, of the average earnings of a household with permanent ability a_i and occupation o_i in the period prior to retiring, working 1/3 of their time.[21] $\psi_{ss,t}$ is set to ensure that the Social Security system breaks even in equilibrium.

3.2 Preferences

The instant utility function, $u(c_{ij}, h_{ij})$, depends on consumption, c_{ij}, and labor supply, $h_{ij} \in (0, 1]$, and is given by:[22]

$$u(c_{ij}, h_{ij}) = \ln c_{ij} - \chi \frac{h_{ij}^{1+\eta}}{1+\eta}, \tag{2}$$

where η is the inverse Frisch elasticity of labor supply. Log utility from consumption ensures the existence of a balanced-growth path for the economy. The utility function of retired households has one extra term, as they gain utility from the bequest they leave to living generations:

$$D(b_{ij+1}) = \varphi \ln(b_{ij+1}), \tag{3}$$

where b_{ij+1} is the level of liquid savings of household i. The expected discounted lifetime utility of household i after occupational choice is given by:

$$V_i = \mathbb{E}_0 \left[\sum_{j=1}^{J} \beta^{j-1} \left[S_j u(c_{ij}, n_{ij}) + (S_j - S_{j+1}) D(b_{ij+1}) \right] \right], \tag{4}$$

where β is the discount factor and $S_j = 1$ for $j \leq 45$.

3.3 Labor Income

Labor productivity depends on four elements that determine the efficiency units of labor each household is endowed with every period: Occupation, o_i, age, j, permanent ability, a_i, and the idiosyncratic productivity shock, u_{ij}, which we assume follows an AR(1) process:

$$u_{ij} = \rho_u u_{ij-1} + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim N\left(-\sigma_{u,o}^2 / [2(1+\rho_u)], \sigma_{u,o}^2\right), \tag{5}$$

where $\sigma_{u,o}$ is the occupation-specific standard deviation of the i.i.d error term, and its mean is set so that $\mathbb{E}[\exp(u_{ij})] = 1$. Thus, household i's hourly wage at age j is given by:

$$w_{it}(j, o_i, a_i, u_{ij}) = w_{ot} \exp\left[\gamma_0 + \gamma_1 j + \gamma_2 j^2 + \gamma_3 j^3 + \vartheta_o a_i + u_{ij}\right], \tag{6}$$

where γ_1, γ_2 and γ_3 is the age profile of wages, and γ_0 is set such that the age polynomial is equal to zero at age $j = 23$, i.e., approximately the mean age of working-age households in the model. w_{ot} is the economy-wide wage rate for that occupation and ϑ_o is the return from

[21] This formulation removes the incentive to work extra hours in the years leading up to retirement to secure a higher retirement benefit, which is contrary to what is observed in the data.

[22] We assume that disutility of work depends only on hours worked, not occupation.

ability in occupation $o \in O \equiv \{\text{NRC}, \text{NRM}, \text{RC}, \text{RM}\}$. ϑ_o is a key parameter in the model, as it determines the extent to which there is self-selection of households into occupations.

3.4 Technology

There exist three competitive final goods: consumption goods, structure investment goods, and equipment investment goods. Each is produced by transforming a single intermediate input using a linear production technology. All payments are made in the consumption good, which is the *numeraire*.

The consumption good is obtained by transforming a quantity $Z_{c,t}$ of intermediate input into output, which is then sold at price $p_{c,t}$ to households and the government. The transformation technology is:

$$C_t + G_t = Z_{c,t}, \tag{7}$$

where $Z_{c,t}$ is the quantity of input, purchased at $p_{z,t}$ from a representative intermediate goods firm. Given that the consumption good is competitively produced, its price equals the marginal cost of production:

$$p_{c,t} = 1 = p_{z,t}. \tag{8}$$

Likewise, structure investment good firms produce output with a similar technology:

$$X_{s,t} = Z_{s,t}, \tag{9}$$

and therefore $p_{s,t} = 1$. The production of $X_{e,t}$, the equipment investment good, uses the transformation technology:

$$X_{e,t} = \frac{Z_{e,t}}{\xi_t}, \tag{10}$$

where $Z_{e,t}$ is the quantity of input z used in the production of the final equipment goods. $1/\xi_t$ is the level of technology used in the production of $X_{e,t}$ relative to the final consumption good. As ξ_t declines, the relative productivity in assembling the equipment good increases. We assume that ξ_t evolves exogenously. We obtain the price of the equipment goods from the zero profit condition:

$$p_{e,t} = \xi_t p_{z,t} = \xi_t, \tag{11}$$

where $\xi_t = p_{e,t}/p_{c,t}$ is interpreted as the relative price of the equipment good.

A representative intermediate goods firm produces $Z_{c,t} + Z_{s,t} + Z_{e,t}$ using a constant returns to scale technology in capital and labor inputs, $Y_t = F(K_{s,t}, K_{e,t}, N_{\text{NRC}t}, N_{\text{NRM}t}, N_{\text{RC}t}, N_{\text{RM}t})$, where $K_{s,t}$ is structure capital and $K_{e,t}$ is capital equipment. The firm rents structure capital at rate $r_{s,t}$, equipment at $r_{e,t}$, and each labor variety at w_{ot}, $o \in O$. Aggregate demand, measured in terms of the consumption good: $Y_t = C_t + G_t + X_{s,t} + \xi_t X_{e,t}$, factor prices, and the price of the intermediate good $p_{z,t}$ are taken as given. The firm chooses capital and labor inputs for each period to maximize profits:

$$\Pi_{z,t} = p_{z,t} Y_t - r_{s,t} K_{s,t} - r_{e,t} K_{et} - \sum_{o \in O} w_{ot} N_{ot}, \tag{12}$$

subject to:

$$Z_{c,t} + Z_{s,t} + Z_{e,t} = C_t + G_t + X_{s,t} + \xi_t X_{e,t} = Y_t. \tag{13}$$

This setup implies that $Z_{c,t} = C_t + G_t$, $Z_{s,t} = X_{s,t}$, $Z_{e,t} = \xi_t X_{e,t}$, and $F(.) = Y_t = C_t + G_t + X_{s,t} + \xi_t X_{e,t}$. We assume the production function of intermediate goods is Cobb-Douglas over structure capital and CES over the remaining inputs:

$$F(.) = A_t P(.) = A_t K_{s,t}^{\alpha} \left[\sum_{l=1}^{3} \varphi_l Z_{l,t}^{\frac{\sigma-1}{\sigma}} + \left(1 - \sum_{l=1}^{3} \varphi_l\right) N_{\text{RM}t}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma(1-\alpha)}{\sigma-1}}, \tag{14}$$

$$Z_{1,t} = \left[\phi_1 K_{e,t}^{\frac{\rho_1-1}{\rho_1}} + (1-\phi_1) N_{\text{NRC}t}^{\frac{\rho_1-1}{\rho_1}} \right]^{\frac{\rho_1}{\rho_1-1}}, \; Z_{2,t} = \left[\phi_2 K_{e,t}^{\frac{\rho_2-1}{\rho_2}} + (1-\phi_2) N_{\text{NRM}t}^{\frac{\rho_2-1}{\rho_2}} \right]^{\frac{\rho_2}{\rho_2-1}},$$

$$Z_{3,t} = \left[\phi_3 K_{e,t}^{\frac{\rho_3-1}{\rho_3}} + (1-\phi_3) N_{\text{RC}t}^{\frac{\rho_3-1}{\rho_3}} \right]^{\frac{\rho_3}{\rho_3-1}},$$

where A_t is total factor productivity, ϕ_l and φ_l are distribution parameters where $l = 1, 2, 3$, correspond to the occupation types NRC, NRM, and RC, respectively.[23] ρ_l is the elasticity of substitution between capital and the nested labor variety i, and σ is the elasticity of substitution between each composite $Z_{l,t}$ and routine manual labor. Complementarity between the two inputs in $Z_{l,t}$ requires that $\rho_l < \sigma$, as in Krusell et al. (2000).

Each variety of labor input is measured in efficiency units, $N_{ot} \equiv h_{ot} \varrho_{ot}$, where h_{ot} is the aggregate amount of labor hours in that occupation and ϱ_{ot} is an efficiency index representing the latent quality per hour worked in occupation type o in period t. ϱ_{ot} can be interpreted as an occupation-specific technology level due to research and development or human capital accumulation. Firm maximization implies that marginal products equal factor prices.[24]

Capital laws of motion are given by:

$$K_{s,t+1} = (1-\delta_s) K_{s,t} + X_{s,t}, \tag{15}$$
$$K_{e,t+1} = (1-\delta_e) K_{e,t} + X_{e,t}, \tag{16}$$

where δ_s and δ_e are the depreciation rates of structures and equipment, respectively.

[23] Krusell et al. (2000), Karabarbounis and Neiman (2014), and Eden and Gaggl (2018) use CES production functions where capital equipment is nested with all labor varieties except for unskilled/routine manual labor, which is introduced in isolation. The reason for this setup is the set of symmetry restrictions on substitution elasticities imposed by the CES functional form, as explained in Krusell et al. (2000). In a nutshell, this nesting form allows for complementarity between capital and differentiated labor (NRC NRM, RC) while permitting the elasticities of substitution between routine manual labor and other labor varieties to be different. Our version extends this framework with a finer breakdown of labor varieties. In estimating the production function, we use the Simulated pseudo-Maximum Likelihood Estimation (SPMLE) method proposed by Ohanian et al. (1997) which was also applied in Krusell et al. (2000). Our application is described in the next section.

[24] The first-order conditions can be found in section C of the Appendix.

3.5 Government

The social security system is managed by the government and runs a balanced budget. The revenues are collected from taxes on employees and on the representative firm at rates τ_{ss} and $\tilde{\tau}_{ss}$, respectively, and are used to pay retirement benefits, Ψ_t.

The government taxes consumption, τ_c, and capital income, τ_k, at flat rates. The labor income tax follows a non-linear functional form as in Benabou (2002), Heathcote et al. (2020) and Holter et al. (2019):

$$y_{a,ij}/AE_t = \theta_0 (y_{ij}/AE_t)^{1-\theta_1}, \tag{17}$$

where θ_0 and θ_1 govern the level and progressivity of the tax schedule, respectively. y_{ij} is the pre-tax labor income and $y_{a,ij}$ is after-tax labor income.[25] We adjust the tax function by Average Earnings (AE) when computing equilibrium so that the tax rate for a person with average earnings in the model continues to equal the tax rate of a person with average earnings in the data.[26]

Tax revenues from consumption, labor, and capital income taxes are used to finance the level of public consumption, G_t, which clears the budget constraint, and the interest, $r_t B_t$, on public debt. Denoting social security revenues by R_t^{ss} and other tax revenues as T_t, the government budget constraint is defined as:

$$T_t = G_t + r_t B_t^G, \tag{18}$$

$$\int_{j>45,o,a} \Psi_t(o,a) d\Phi_t = R_t^{ss}. \tag{19}$$

3.6 Asset Structure

Households hold three asset types: risk-free government bonds, b_{ij}^G, structures capital, $k_{s,ij}$, and equipment capital, $k_{e,ij}$. We assume a no-arbitrage condition between all assets. Thus, the return rates must satisfy:

$$\frac{1}{\xi_t}\left[\xi_{t+1} + (r_{e,t+1} - \delta_e \xi_{t+1})(1-\tau_k)\right] = 1 + (r_{s,t+1} - \delta_s)(1-\tau_k), \tag{20}$$

and

$$1 + r_{t+1}(1-\tau_k) = 1 + (r_{s,t+1} - \delta_s)(1-\tau_k). \tag{21}$$

The purchase value of the household portfolio at the end of period t is defined as:

$$b_{ij+1} \equiv \xi_t k_{e,ij+1} + k_{s,ij+1} + b_{ij+1}^G. \tag{22}$$

3.7 Household Problem

In any given period a working-age household is defined by its age, j, occupation o_i, asset position b_{ij}, permanent ability a_i, and persistent idiosyncratic productivity shock u_{ij}. After

[25] See the Appendix of Holter et al. (2019) for a detailed discussion of the properties of this tax function.

[26] Since the tax function is not scale invariant, the same normalization is necessary when estimating and comparing the tax system for different years in the data.

selecting an occupation, the household chooses consumption, c_{ij}, work hours, h_{ij}, and future asset holdings, b_{ij+1}, to solve its problem of maximizing expected utility.[27] The household budget constraint is:

$$c_{ij}(1+\tau_c)+\xi_t k_{e,ij+1}+k_{s,ij+1}+b^G_{ij+1}\leq \left[\xi_t+(r_{e,t}-\delta_e\xi_t)(1-\tau_k)\right]k_{e,ij}$$
$$+\left[1+(r_{s,t}-\delta_s)(1-\tau_k)\right]k_{s,ij}+\left[1+r_t(1-\tau_k)\right]b^G_{ij}+\Gamma_t\left[1+r_t(1-\tau_k)\right]+Y^N, \tag{23}$$

where Γ_t is the bequest received per capita, and Y^N is the household's labor income after social security and labor income taxes. In equilibrium, the budget constraint can be rewritten, by using (20) and (21), as:

$$c_{ij}(1+\tau_c)+b_{ij+1}\leq (b_{ij}+\Gamma_t)\left[1+r_t(1-\tau_k)\right]+Y^N. \tag{24}$$

The household problem can be formulated recursively as:

$$V(j,b_{ij},o_i,a_i,u_{ij})\leq \max_{c_{ij},h_{ij},b_{ij+1}}\left[u\left(c_{ij},h_{ij}\right)+\beta\mathbb{E}_{u_{j+1}}\left[V(j+1,b_{ij+1},o_i,a_i,u_{ij+1})\right]\right],$$

s.t.:

$$c_{ij}(1+\tau_c)+b_{ij+1}=(b_{ij}+\Gamma_t)\left[1+r_t(1-\tau_k)\right]+Y^N,$$
$$Y^N=\frac{h_{ij}w_{it}\left(j,o_i,a_i,u_{ij}\right)}{1+\tilde{\tau}_{ss}}\left(1-\tau_{ss}-\tau_l\left[\frac{h_{ij}w_{it}\left(j,o_i,a_i,u_{ij}\right)}{1+\tilde{\tau}_{ss}}\right]\right),$$
$$h_{ij}\in(0,1],\quad b_{ij}\geq 0,\quad b_{i0}=0\quad\forall i,\quad c_{ij}>0.$$

The problem of a retired household differs in three ways: There is a positive age-dependent probability of dying, $\pi(j)$, a bequest motive $D(b_{ij+1})$, and labor income is replaced by a constant retirement benefit depending on permanent ability, $\Psi_t(a_i,o_i)$. The retired household's problem can be written as:

$$V(j,b_{ij},o_i,a_i)=\max_{c_{ij},b_{ij+1}}\left[u\left(c_{ij},b_{ij+1}\right)+\beta(1-\pi(j))V(j+1,b_{ij+1},o_i,a_i)+\pi(j)D(b_{ij+1})\right],$$

s.t.:

$$c_{ij}(1+\tau_c)+b_{ij+1}\leq (b_{ij}+\Gamma_t)\left[1+r_t(1-\tau_k)\right]+\Psi_t(a_i,o_i),$$
$$b_{ij+1}\geq 0,\quad c_{ij}>0.$$

3.8 Recursive Competitive Equilibrium

In the stationary recursive competitive equilibrium, which we use to characterize the U.S. economy in 1980, agents optimize, given prices and budget constraints, markets clear, prices

[27] For simplicity, we omit the t subscript from the value function and the household policy functions, using it only for prices, aggregate quantities, and distributions. However, all these objects depend on time to the extent that aggregate quantities and prices change.

are determined by their marginal products, the government budget constraint balances, and the cross-sectional distribution across household types is stationary. For sake of brevity, the formal equilibrium definition is stated in Appendix E.1. Appendix F.2 describes the algorithm to compute it.

The main quantitative and the optimal policy analyses are performed by introducing unexpected changes to technology or fiscal parameters (i.e., an MIT shock). This generates a deterministic path for the economy, which converges to a long-run stationary recursive competitive equilibrium with the new parameter values. The definition of an equilibrium along the transition path is similar to the stationary equilibrium described in Appendix E.1, but for the fact that distributions, aggregate quantities, prices, and policy and value functions are indexed by time. Appendix E.2 formally defines the transition equilibrium and Appendix F.2 describes the computational algorithm.

4 Estimating the Production Function

In this section, we describe the stochastic specification of the production function model, the equations to be estimated, and the results. The estimation strategy follows Krusell et al. (2000). When we calibrate our model, we will treat the parameter estimates from this section as exogenously given. When we study the impact of changes in technology over time on inequality, we will insert our results from this section in the model.[28]

4.1 Stochastic Specification

The stochastic elements in our model are the unobserved technology components: (i) the relative technological level of the investment good sector; (ii) the set of labor-specific efficiency indices; and (iii) the factor-neutral technological process. We assume that the relative price of equipment ($\tilde{\xi}_t = \xi_t/\xi_{t-1}$) is trend stationary, and confirm this with a Dickey-Fuller test. We assume that the labor efficiency index processes have different linear trends for each labor variety. Defining the processes in logs, we have:

$$\psi_t \equiv \ln(\varrho_t), \quad \psi_t = \psi_0 + \psi_1 t + \nu_t, \tag{25}$$

where ψ_t is a (4×1) vector of the log of the latent efficiency indices, ψ_0 is a (4×1) vector of constants which specify the value of the indices at the beginning of the sample, ψ_1 is a (4×1) vector of growth rates, and ν_t is a (4×1) vector of shock processes that we assume to be multivariate normal, i.i.d. with covariance matrix Ω: $\nu_t \sim N(0, \Omega)$. The i.i.d. assumption simplifies the identification of the factor-neutral technological change, A_t, which is described below.

[28] The data used in the estimation is described in Appendix B.

4.2 Equation Specification

We use a system with two sets of equations obtained from the first order conditions of agents to estimate the model: (i) the wage bills relative to the routine manual labor variety; and (ii) a no-arbitrage condition between investing in equipment and structure capital. These are defined as follows:

$$\frac{w_{o,t}h_{o,t}}{w_{\mathrm{RM},t}h_{\mathrm{RM},t}} = wbr_{o,t}(\psi_t, X_t; \theta), \qquad o \in O = \{\mathrm{NRC}, \mathrm{NRM}, \mathrm{RC}\}, \tag{26}$$

and

$$1 + [F_{K_s}(\psi_{t+1}, X_{t+1}; \theta) - \delta_{s,t+1}] = E_t\left(\frac{\xi_{t+1}}{\xi_t}\right)(1 - \delta_{e,t+1}) + \frac{F_{K_e}(\psi_{t+1}, X_{t+1}; \theta)}{\xi_t}, \tag{27}$$

where equation (27) is obtained from equation (20), assuming that $\xi_t \neq \xi_{t+1}$, and where we substituted the return rates by factor marginal productivities.[29]

Depreciation rates are indexed by t since they change over the time.[30] The relative wage bills in the model $wbr_{o,t}$ are functions of X_t and θ. X_t is the vector of inputs and depreciation rates $\{K_{s,t}, K_{e,t}, h_{\mathrm{NRC}t}, h_{\mathrm{NRM}t}, h_{\mathrm{RC}t}, h_{\mathrm{RM}t}, \delta_{s,t}, \delta_{e,t}\}$. The vector θ is the set of parameters $\{\alpha, \rho_1, \rho_2, \rho_3, \phi_1, \phi_2, \phi_3, \varphi_1, \varphi_2, \varphi_3, \psi_0, \psi_1, \Omega, \eta_\omega, K_{e,0}\}$, including the first observation of the equipment capital stock, which we estimate jointly with the other parameters. η_ω is the standard deviation of the error term in the equipment price equation, which we specify below. Like Krusell et al. (2000), we assume that there is no risk premium in equation (27), and that the tax treatment is identical between equipment and structure capital returns. Finally, we substitute the first term on the right hand side of equation (27) with $E_t\left(\xi_{t+1}/\xi_t\right)(1 - \delta_{e,t}[1 - \tau_{k,t}]) + \omega_t$, where ω_t is the i.i.d. forecast error and $\omega_t \sim N\left(0, \eta_\omega^2\right)$. This set of assumptions imply that $A_t = Y_t/P(.)$ from equation (14).

Given that this is a non-linear system of eight equations with unobserved state variables, standard linear Kalman filter techniques cannot be applied to estimate the parameter vector θ. Ohanian et al. (1997) propose a two-step version of the SPML estimator to find θ for this type of problem.[31]

The parameter vector θ has dimension 36. Our sample contains 49 observations for each equation. We reduce the number of parameters estimated by external calibration or by setting *a priori* restrictions. First, we impose that Ω be a diagonal matrix and that the variance of the disturbances is identical for all labor types. Thus, $\Omega = \eta_v^2 I_4$, where η_v^2 is the common innovation variance and I_4 is a (4×4) identity matrix. Second, we fix $\psi_{4,0}$, the initial level of the latent efficiency index of routine manual workers, which is not identified. Third, we set the income share of structures to 0.04 as in Krusell et al. (2000). Finally, we regress the variation

[29] Note that this no-arbitrage equation applies on capital returns net of depreciation. Hence, in equilibrium, we are allowing for different capital gross returns across the two types of capital because they have different depreciation rates.

[30] See Appendix B for the method of construction of the depreciation rates.

[31] See Appendix D for a detailed explanation of our application.

rate of the relative price of equipment on a linear trend to calibrate the forecast error variance of the equipment price index. We set η_ω to be equal to the estimated standard deviation of the error term in the regression $\tilde{\sigma}_\omega = 0.032$. This reduces the number of parameters to be estimated to 19: The common variance of the latent processes, η_v^2, the elasticities, $\sigma, \rho_1, \rho_2, \rho_3$, the production function share parameters, $\phi_1, \phi_2, \phi_3, \varphi_1, \varphi_2, \varphi_3$, the parameters governing the latent state variables, except for $\psi_{4,0}$, and the initial level of capital equipment, $K_{e,0}$.

4.3 Estimation Results and Model Fit

The model is estimated using data from 1967 to 2016 and the Simulated Pseudo Maximum Likelihood Estimation (SPMLE) procedure. Table 1 shows the resulting estimates.

Elasticity estimates for the nested occupation types are all consistent with capital-occupation complementarity, i.e., $\sigma > \rho_i$, $i = 1, 2, 3$. The estimation of these elasticities is one of the contributions of this paper to the literature.

The most comparable estimates are provided by Eden and Gaggl (2018), who specify a CES production function with non-routine labor nested with capital. In contrast to our estimates of 0.5 and 2.1 for NRC and NRM labor, they estimate an elasticity of substitution of 1.4 for non-routine labor. For routine manual labor, their estimate is 8.0 for routine occupations, compared to our elasticity of 5.6 for RM. Although less comparable, Krusell et al. (2000) obtain a value of 0.67 for skilled labor and 1.67 for unskilled labor. For the processes of occupation-specific technology, we estimate that only the non-routine cognitive occupations have experienced positive growth. At the same time, routine manual labor has suffered the largest decline.[32]

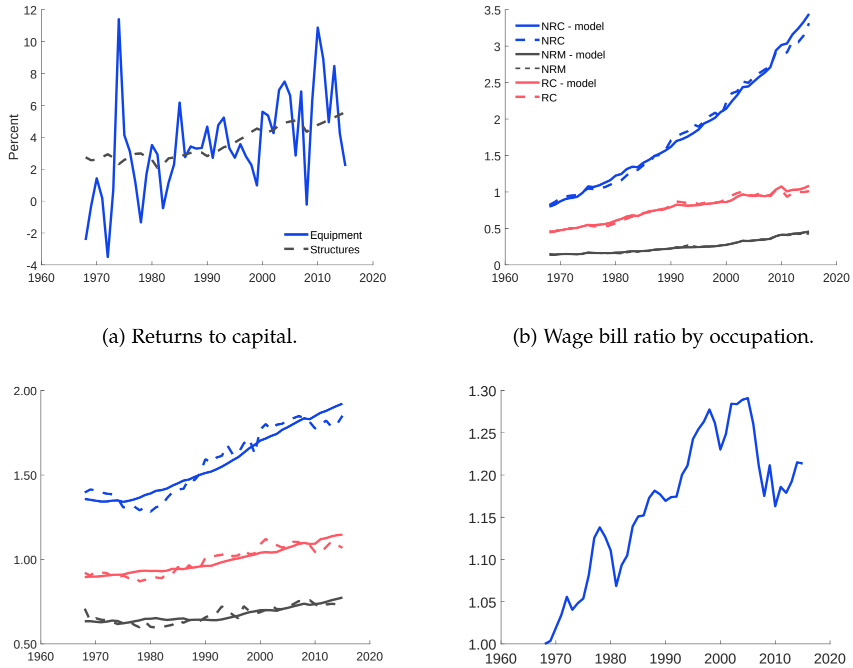
Table 1: Parameter Estimates

Parameter	Description	Value	Parameter	Description	Value
σ	EOS RM	5.564	ρ_1	EOS NRC	0.497
ρ_2	EOS NRM	2.055	ρ_3	EOS RC	5.029
ϕ_1	Share NRC	0.378	ϕ_2	Share NRM	0.086
ϕ_3	Share RM	0.279	φ_1	Share comp. NRC	0.160
φ_2	Share comp. NRM	0.045	φ_3	Share comp. RC	0.023
$\psi_{0,1}$	Intercept NRC	0.859	$\psi_{0,2}$	Intercept NRM	1.936
$\psi_{0,3}$	Intercept RC	3.582	$\psi_{1,1}$	Slope NRC	0.002
$\psi_{1,2}$	Slope NRM	−0.006	$\psi_{1,3}$	Slope RC	−0.001
$\psi_{1,4}$	Slope RM	−0.010	$K_{e,0}$	Init. equip. capital	582

Note: The table shows the parameter estimates for the production function and the labor efficiency indices. "EOS" stands for elasticity of substitution. The ϕ are the shares of each occupation inside each labor-equipment composite. The φ are the shares of each labor-equipment composite. The ψ_0 indicate the intercept of the linear labor efficiency indices, and ψ_1 the slope. $K_{e,0}$ is the starting level of equipment capital in millions of dollars.

Figure 2 shows model fit to targeted moments over time. Figure 2a displays aggregate ex-post return rates of equipment and structures implied by our model, the difference being

[32] Vom Lehn (2020) also estimates elasticities of substitution between different task bundles that are not directly comparable to ours. In his production function, abstract and manual labor inputs are substitutes or complement to a bundle composed of routine labor input and capital equipment. In contrast, in our framework, NRC, NRM and RC all have a constant elasticity of substitution with capital equipment directly. In the case of homogeneous workers, he calibrates the elasticity of substitution between routine labor input and capital equipment to 1.3, between manual and a bundle of routine labor input and equipment to 1.49, and between abstract labor input and a bundle of routine labor input and equipment to 0.31.



(a) Returns to capital.

(b) Wage bill ratio by occupation.

(c) Wage premia.

(d) Total factor productivity.

Note: The model predictions of the variables presented in this figure are computed based on the production functions parameters and observed data on the production inputs from 1968 to 2015—we lose both the first and the last period of the sample to estimate the model. In Figure 2d, total factor productivity is normalized to 1 in 1968. Construction of the measures is described in Appendix B.

Figure 2: Empirical Model Fit to Targeted and Non-Targeted Moments.

zero in expectation as dictated by the no-arbitrage condition. They have a 4% average, as in Krusell et al. (2000), although a slightly increasing trend from the early 2000s onward.

Figure 2b plots wage bill ratios implied by the model, as specified by the set of equations (26), and the data. Model predictions closely track the data. The NRC wage bill shot up from near par with routine manual labor in 1968 to 3.5 in 2015. In contrast, NRM and RC wage bills grow slowly upwards relative to that of routine manual occupations, which is explained by both their lower level of complementarity with equipment capital as well as their declining level of latent efficiency.

Figure 2c shows the model fit to the wage premia of each occupation relative to RM. As in the previous figure, the dashed lines indicate the data and the solid lines are the model predictions. In all cases, the model tracks the data closely. This is important given that our goal is to use the estimated parameters to calibrate the theoretical model. The key force driving earnings dispersion is the change in wage premia across groups.

Finally, Figure 2d displays our estimate of total factor productivity in the U.S. for this period. From 1968 to 2008, TFP increased by almost 30% and then fell to around 20% in the following years. For comparison, the estimate of total factor productivity by the Penn World Table increases by 30% from 1968 to 2015 (FRED).

4.4 Discussion of Automation and Our Production Structure

Our production structure is designed to flexibly capture empirically relevant substitution patterns between capital and labor inputs. In particular, RM labor is embedded in a CES composite that also includes both equipment capital and NRC labor. This setup allows for general substitution possibilities but implies that RM labor is equally substitutable with capital and NRC labor.

Alternative perspectives in the literature adopt different stances on how automation affects capital-labor interactions. For example, Acemoglu and Restrepo (2022) propose a framework in which automation technologies act as task-specific capital that substitutes directly for routine labor, leading to displacement effects and employment polarization. In contrast, Aghion et al. (2022) and Aghion and Jaravel (2023) argue that modern capital, particularly in manufacturing, can exhibit broad complementarities with both high-skill and RM labor. Their findings suggest that capital adoption need not reduce demand for routine workers and may in fact support broader labor demand growth.

While a production function more closely aligned with automation-focused models — such as nesting RM labor directly with capital—could generate sharper substitution effects, the implications are not straightforward in our framework. Isolating RM from the RM–NRC–capital nest could reduce complementarity between NRC labor and capital, a margin that plays an important role in our results. Furthermore, our model includes a latent occupational productivity process that is also estimated to match wage bill ratios. The interaction between these productivity estimates and alternative aggregator structures makes the net effect of respecifying the nesting structure theoretically ambiguous.

Moreover, because our estimation is based on macroeconomic data at the occupational level, the identification of automation-specific mechanisms is limited. Structural modeling of automation—particularly at the task level—would require richer micro-level data that go beyond the scope of this paper. We therefore view the exploration of alternative aggregator forms and explicit modeling of automation technologies as promising avenues for future research.

In conclusion, we provide new estimates for the elasticities of substitution between equipment capital and the occupation categories defined in Autor et al. (2003). We find that our model is broadly compatible with the data, especially with respect to the occupation wage premia, which is crucial for ensuring that the predictions of the theoretical model are consistent with the data. We now turn to the calibration of the theoretical model, which uses the estimates obtained from this section to parameterize the production side of the economy.

5 Calibration

This section describes the calibration of the benchmark model to resemble the U.S. economy in 1980. Many parameters are set externally (i.e. we estimate them directly from the data or take them from the literature). This includes the production function parameters we estimated using the procedure described in Section 4 but also the tax function and the age profile of earnings. Table 2 lists the externally calibrated parameter values and data sources. The fifteen parameters in Table 3 are estimated by the simulated method of moments (SMM) approach, where we find parameter values that minimize the distance between selected model and data moments.

5.1 Externally Calibrated Parameters

Below we discuss the external calibration of parameters that were not estimated using the procedure described in Section 4. Table 2 summarizes the externally calibrated parameters.

Table 2: Externally Calibrated Parameters in the Benchmark Economy

Description	Parameter	Value	Source
Preferences			
Inverse Frisch elasticity	η	3.000	Literature
Labor productivity			
Parameter 1 age profile of wages	γ_1	0.265	Brinca et al. (2016)
Parameter 2 age profile of wages	γ_2	-0.005	Brinca et al. (2016)
Parameter 3 age profile of wages	γ_3	3.6×10^{-5}	Brinca et al. (2016)
Persistence of transitory shock	ρ_u	0.335	Brinca et al. (2016)
Technology			
Equipment depreciation rate	δ_e	0.106	Section 4
Structures depreciation rate	δ_s	0.026	Section 4
Share structures	α	0.040	Section 4
Share NRC	ϕ_1	0.378	Section 4
Share NRM	ϕ_2	0.086	Section 4
Share RC	ϕ_3	0.279	Section 4
Share composite NRC	φ_1	0.160	Section 4
Share composite NRM	φ_2	0.045	Section 4
Share composite RC	φ_3	0.023	Section 4
EOS NRC	ρ_1	0.497	Section 4
EOS NRM	ρ_2	2.055	Section 4
EOS RC	ρ_3	5.029	Section 4
EOS RM	σ	5.564	Section 4
Latent efficiency NRC	ϱ_1	2.734	Section 4
Latent efficiency NRM	ϱ_2	4.955	Section 4
Latent efficiency RC	ϱ_3	34.662	Section 4
Latent efficiency RM	ϱ_4	0.378	Section 4
Total factor productivity	A	16.728	Section 4
Relative price of investment goods	ξ	1.000	Normalization
Government and SS			
Consumption tax rate	τ_c	0.054	Mendoza et al. (1994)
Capital income tax rate	τ_k	0.469	Mendoza et al. (1994)
Tax scale parameter	θ_0	0.850	Wu (2020)
Tax progressivity parameter	θ_1	0.187	Wu (2020)
SS tax employees	τ_{ss}	0.061	Social Security Bulletin, July 1981
SS tax employers	$\tilde{\tau}_{ss}$	0.061	Social Security Bulletin, July 1981
Debt to GDP	B^G/Y	0.320	FRED

Preferences We set the inverse of the Frisch elasticity of labor supply, η, to 3, which is a standard value in the literature.

Labor productivity The wage profile through the life cycle (see equation 6) is calibrated using estimates for the United States from Brinca et al. (2016). However, instead of using a common error variance for the transitory shock, we use occupation-specific error variances to target data moments. We adopt this approach for two reasons. First, to guarantee the variance of earnings inequality in the model and in the data have the same starting point, so that comparisons are straightforward. Second, so we can use the occupation-specific error variances as an instrument to fit observed earnings inequality exactly in 2015 to carry out optimal policy experiments in section 7.

Technology Equipment and structure depreciation rates are set to match those used in the estimation of the empirical model for 1980, and described in Appendix B. The production function is calibrated using the parameters estimated from the empirical model. The efficiency indices for each occupation are set to match those of the empirical model in 1980. The level of total factor productivity is set to the estimate from the empirical model for 1980.

Government We set θ_0 and θ_1 to the estimates obtained by Wu (2021) for 1980. For social security rates, we assume no progressivity. Both social security tax rates, employer and employee, are set to 0.06, the average rate in 1980. We set τ_c and τ_k to match the values obtained in Mendoza et al. (1994) for 1980, i.e. $\tau_c = 0.05$, $\tau_k = 0.47$. Finally, we set government debt-to-GDP B^G/Y to 0.32, the value observed in 1980 from FRED.

5.2 Endogenously Calibrated Parameters

To calibrate the remaining parameters, $\{\beta, \chi, \varphi, \vartheta_{\text{NRC}}, \vartheta_{\text{NRM}}, \vartheta_{\text{RC}}, \vartheta_{\text{RM}}, \sigma_{u,\text{NRC}}, \sigma_{u,\text{NRM}}, \sigma_{u,\text{RC}}, \sigma_{u,\text{RM}}, \mu_{\kappa,\text{NRC}}, \mu_{\kappa,\text{NRM}}, \mu_{\kappa,\text{RC}}\}$, we use a simulated method of moments approach, for which we construct the following loss function:

$$L(\tilde{\theta}) = ||M_m - M_d||, \tag{28}$$

where $\tilde{\theta}$ is the vector of parameters to be estimated, and M_m and M_d the moments in the model and in the data for 1980, respectively. Our estimate, $\tilde{\theta}^*$, is obtained by minimizing (28).

Table 3: Parameters Calibrated Internally.

Parameter	Value	Description
φ	4.45	Bequest utility
β	0.96	Discount factor
χ	65.14	Disutility of work
ϑ_{NRC}	0.37	Return to ability (NRC)
ϑ_{NRM}	0.12	Return to ability (NRM)
ϑ_{RC}	0.47	Return to ability (RC)
ϑ_{RM}	0.22	Return to ability (RM)
$\sigma_{u,\text{NRC}}$	0.45	Transitory shock s.d. (NRC)
$\sigma_{u,\text{NRM}}$	0.52	Transitory shock s.d. (NRM)
$\sigma_{u,\text{RC}}$	0.40	Transitory shock s.d. (RC)
$\sigma_{u,\text{RM}}$	0.41	Transitory shock s.d. (RM)
$\mu_{\kappa,\text{NRC}}$	-5.73	Location - taste shock (NRC)
$\mu_{\kappa,\text{NRM}}$	2.86	Location - taste shock (NRM)
$\mu_{\kappa,\text{RC}}$	0.05	Location - taste shock (RC)
σ_κ	5.00	Common scale - taste shock

Note: All parameters are calibrated internally to match model moments. See Section 5 for details.

Table 3 presents the parameters calibrated internally through SMM estimation and their

Table 4: Model Fit to Data in 1980.

Data Moment	Source	Value	Model Value
Avg. wealth retirees to avg. wealth	US Census Bureau	1.31	1.31
Capital to output	BEA and CPS	1.41	1.40
Fraction of hours worked	BEA	1/3	1/3
NRC wage premium	CPS	1.28	1.31
NRM wage premium	CPS	0.60	0.62
RC wage premium	CPS	0.88	0.87
Variance of log earnings	CPS	0.45	0.44
Variance of log earnings (NRC)	CPS	0.41	0.41
Variance of log earnings (NRM)	CPS	0.41	0.41
Variance of log earnings (RC)	CPS	0.41	0.41
Variance of log earnings (RM)	CPS	0.30	0.30
Employment share (NRC)	CPS	0.31	0.31
Employment share (NRM)	CPS	0.10	0.10
Employment share (RC)	CPS	0.24	0.24

Note: Employment shares may not add up to one due to rounding.

estimates. We use the ratio between the average wealth of 65 and older to the average wealth in the economy as the target for the utility of bequests parameter. The discount factor is set by targeting the capital-to-output ratio.[33] The capital stock is obtained from the estimation of the empirical model of Section 4. Disutility from work targets average hours worked.

The parameters governing the wage process and occupational choice jointly determine key inequality statistics in the model. The returns to ability by occupation affect both wage premia and the overall variance of log earnings. This occurs because the average productivity endowment in each occupation is a function of the permanent ability distribution of the individuals selecting into it. These returns influence wage premia through the extensive-margin response of occupational labor supply, and they shape total earnings inequality via both within-occupation ability dispersion and between-occupation differences.

The within-group variance of log earnings is disciplined through the occupation-specific transitory shock variances. The location parameters and the scale parameter of the taste shock jointly govern the occupational employment shares of the NRC, NRM, and RC workers. We find that the set of data moments used in estimation is highly informative and delivers strong identification of these parameters.

In particular, the scale parameter is sharply identified. Lowering the scale from its baseline value of 5 to 3 increases the objective function to 0.164, while raising it to 8 yields a value of 0.103. In contrast, the baseline value achieves a much lower distance of 0.055. These results suggest that empirical moments provide strong discipline for the scale parameter.[34]

Table 4 displays the fit of the model moments to the data moments. In general, no model moment has more than a 5% difference from its empirical target.

[33] This quantity is not directly comparable to the usual K/Y in the macro literature because we include non-residential capital in the form of non-residential structures and equipment capital only, as proposed by Krusell et al. (2000).

[34] Appendix Section G further validates this choice by showing that the implied occupational elasticities to changes in tax progressivity are broadly consistent with the data.

6 Technological Change and Earnings Inequality

We now use the benchmark model to answer the first question posed in the introduction: To what extent does technological change explain the observed increase in earnings inequality? In Section 7 we turn to the policy implications of a technology driven increase in inequality.

6.1 The Sources of Growing Earnings Inequality

Our main quantitative experiment is a 100-year transition starting in 1980, where we let the technology and tax parameters gradually change over time, moving from their 1980 to their 2015 values, while holding all other model parameters fixed.[35] The transition is implemented as an MIT shock, with parameters evolving linearly over time. In other words, in 1980 the agents in our economy suddenly learn of the new path for prices and tax policies over the next 100 years. We simulate the transition path of the economy and measure the resulting change in the variance of log earnings and wage premia 35 years after the shock—that is, in 2015. We do not attempt to forecast future technology or tax changes and assume that all parameters stay constant after 2015, although the simulated transition continues until 2080.

We compare the simulated change in earnings inequality to its empirical counterpart and assess how much of it can be attributed to three channels: Investment-specific technological change (ISTC), labor augmenting technological change (LAT), and TFP growth. We also evaluate the role of tax policy changes over the same period, including the observed decline in capital taxation and shifts in labor income tax progressivity. Each channel is isolated by varying one component at a time, starting from the 1980 steady state.

All other structural parameters — preferences, individual productivity dynamics, and the production function — are held fixed throughout the transition. In particular, the age profile of wages (γ_1, γ_2, γ_3), the idiosyncratic productivity process (ρ_u, $\sigma_{u,o}$), preference parameters (λ, η, β, σ_κ, $\mu_{\kappa,o}$ for $o \in O$), the returns to ability across occupations ($\sigma_{a,o}$ for $o \in O$), and the production function's elasticities and factor shares remain unchanged.

The parameters that vary between 1980 and 2015 are listed in Table 5. We set the relative price of investment goods in 2015 (and later years) to 41% of its 1980 level, consistent with the decline observed in the data between 1980 and 2015. The labor efficiency indices change to their 2015 values using the functional forms estimated in Section 4, and TFP is also updated gradually to its 2015 level. The level and progressivity of the labor income tax schedule are set to the 2015 estimates of Wu (2021). Social Security contributions follow the statutory rates described in Brinca et al. (2016), and effective consumption and capital income taxes are computed using the method of Mendoza et al. (1994) for 2015. Depreciation rates and government debt are also updated using the same method described in Section 5.

[35] There are many other factors that potentially changed between 1980 and 2015, and that could cause either higher or lower inequality (for example, aging population and shrinking gender wage gap, see Wu, 2021) but we focus only on fiscal parameters and technology. For the optimal policy exercises, we also consider a scenario where the variances of the innovations to the idiosyncratic productivity shocks are used to match the changes in within-occupation variance of earnings, see Section 7.3.

Table 5: Parameter Changes 1980-2015.

Parameter	Description	1980	New SS
τ_c	Consumption tax	0.054	0.050
τ_k	Capital income tax	0.469	0.360
τ_{ss}	Employee SS tax	0.061	0.077
$\tilde{\tau}_{ss}$	Employer SS tax	0.061	0.077
δ_e	Equipment dep. rate	0.106	0.148
δ_s	Structures dep. rate	0.026	0.031
B^G/Y	Gov. debt to GDP	0.320	1.020
θ_0	Tax scale	0.850	0.922
θ_1	Tax progressivity	0.187	0.137
ξ	Relative eq. investment price	1.000	0.405
A	TFP	16.728	18.281
ϱ_1	Latent efficiency NRC	2.734	2.986
ϱ_2	Latent efficiency NRM	4.955	4.051
ϱ_3	Latent efficiency RC	34.662	33.907
ϱ_4	Latent efficiency RM	0.378	0.267

Table 6: Model Fit and Projections.

	1980		2015	
Variable	Model	Data	Model	Data
Employment shares				
NRC	0.31	0.31	0.42	0.42
NRM	0.11	0.10	0.11	0.14
RC	0.24	0.24	0.24	0.22
RM	0.35	0.35	0.24	0.22
Hours worked				
Total	0.33	0.33	0.34	0.34
NRC	0.33	0.35	0.34	0.35
NRM	0.33	0.28	0.33	0.30
RC	0.33	0.31	0.34	0.33
RM	0.33	0.35	0.34	0.36
Wage premia				
NRC	1.32	1.28	1.78	1.85
NRM	0.62	0.60	0.74	0.72
RC	0.87	0.88	1.06	1.07
Variance of log earnings				
Total	0.44	0.45	0.53	0.57
NRC	0.41	0.41	0.43	0.51
NRM	0.41	0.41	0.43	0.37
RC	0.41	0.41	0.43	0.49
RM	0.30	0.30	0.31	0.39

Note: The empirical counterpart of the model wage premium is described in Appendix B.

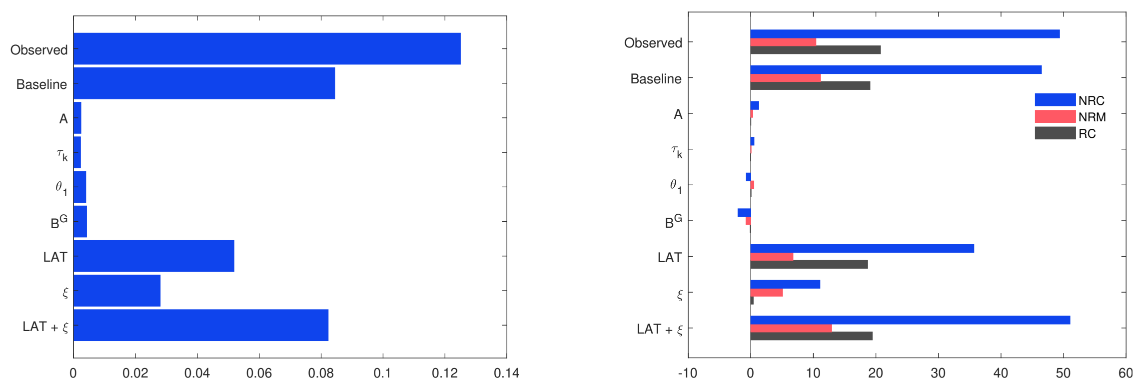
Table 6 compares the model and empirical moments in 1980 and 2015, when we simultaneously change all the parameters of interest from their 1980 values to their 2015 values along the transition. Note that the model projections for 2015 are non-targeted, in the sense that they result from household choices (e.g., occupation choice and saving) in response to parameter changes.

The first panel compares relative input quantities in the theoretical model to those implied by the empirical production function estimates in Section 4. Most moments are well matched, though the model significantly understates the growth in equipment capital to 2015, reflecting its inability to fully replicate the observed rise in aggregate capital.[36]

The second and third panels show employment shares and wage growth by occupation. The model correctly projects the change in NRC and RM employment shares, and the change rate in wages. However, it slightly under-predicts the extent of wage growth and generates no change in employment shares for the NRM and RC between 1980 and 2015.

The fourth and fifth panels display occupation wage premia relative to RM occupations and the variance of log earnings, both total and within occupations. The model predicts the change in wage premia very closely, with all occupations experiencing an increase in the market wage rate relative to RM occupations. Note that this does not follow directly from the model calibration. It is also the result of savings behavior and occupational choices, both of which were calibrated to match empirical moments in 1980, but otherwise respond only to the changing quantities and prices that emerge from model mechanisms in response to technological and tax changes.

Together with a shift in employment shares toward occupations with greater earnings inequality, the changes in wage premia lead to an increase in the variance of log-earnings from 1980 to 2015, as in the data. Combined, the mechanisms in our model account for 67% of the increase in the variance of log earnings. As can be observed, our mechanisms mostly affect between-occupation inequality.[37]



(a) Change in the variance of log pre-tax earnings. (b) Change in wage premia (p.p.).

Note: The bar denoted "Observed" indicates the change in the indicator recorded in the data between 1980 and 2015. The data used is from the CPS and is described in section A of the Appendix. "Baseline" indicates the change predicted in the theoretical model from 1980 to 2015. Each of the remaining bars indicates the change in the model statistics resulting from keeping the corresponding parameters at their 1980 levels. "A" is the change in total factor productivity. τ_k is is the change in capital income tax. θ_1 is the change in progressivity. B^G is the change in government debt to GDP. ξ is investment-specific technological change reflected in the change in the relative price of equipment investment. "LAT" is the change in the set of occupation-specific efficiency indices. "LAT + ξ" is both labor-augmenting technological change and ISTC.

Figure 3: Model Decomposition of the Change in Earnings Inequality from 1980 to 2015.

To understand the drivers of the change in aggregate earnings inequality and to answer

[36] A likely explanation is that the U.S., as a large open economy, experienced substantial capital inflows over the period. According to BEA data, the stock of foreign direct investment (FDI) in the U.S. rose from 83\$ billion in 1980 to 3.4\$ trillion in 2015—a forty-fold increase. See also Chakraborty et al. (2017) on the expansion of cross-border lending to U.S. firms.

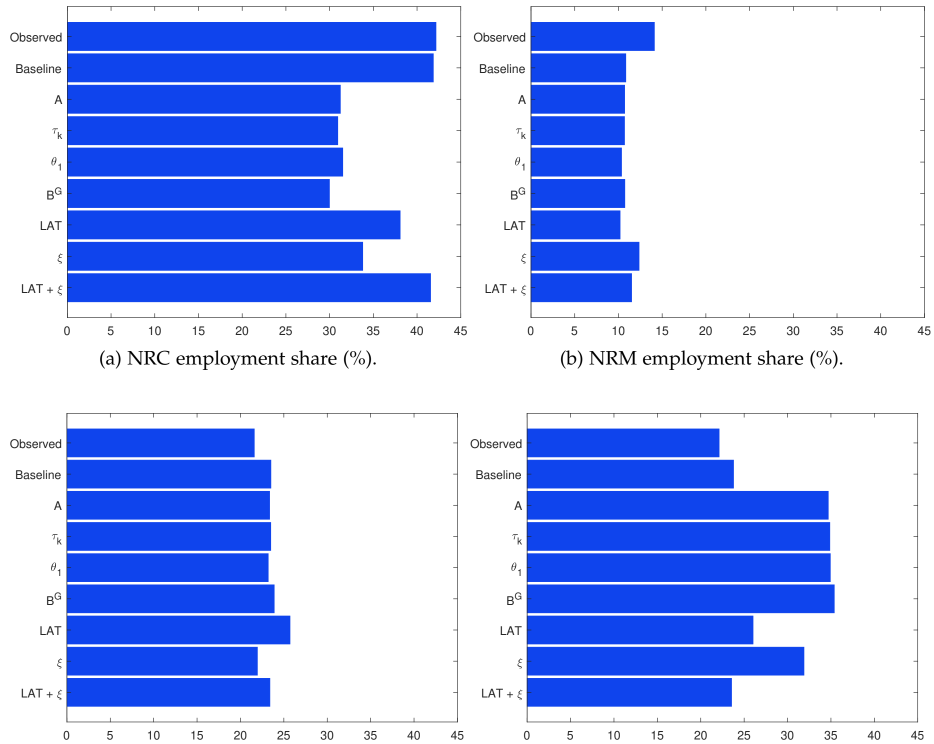
[37] Within-occupation inequality also changes slightly in the model, but not sufficiently to match observed changes which are driven by mechanisms not captured by the model. In Section 7 we consider a scenario where $\sigma_{u,o}$ changes during the transition to explain the changes in within-occupation inequality.

the first question of this paper, we use the model to generate counterfactuals. This is done by starting from the 1980 steady state and changing each set of parameters of interest to their 2015 levels during the transition while keeping the remaining parameters at their 1980 levels. We then compare the resulting change to the variation observed in the data. This allows us to isolate the effects of changing those parameters only, while accounting for the behavioral responses of households and any potential general equilibrium effects.

Figure 3 shows the results of these exercises by displaying the response of labor earnings dispersion measures to a selection of the parameter shifts presented in Table 5. The bars labeled "Observed" in each panel indicate the observed change in that measure. The other bars indicate the change predicted by the model due to shifts in those parameters only. The bars labeled "Baseline", are the model predictions when all the parameters in Table 5 change linearly to their 2015 values.

Figure 3a shows how the model fares in generating a shift in pre-tax log earnings variance compared to the data. The baseline projection of the change in log earnings variance is 67% of the one observed in the data. We find that the main driver of the model's prediction is latent occupation-biased technological change, which generates a 41% increase in the variance of log earnings. According to our estimates, LAT has a decreasing trend for all occupations except for NRC. In particular, the wage rate for RM drops sharply. In response to the increase in the NRC wage premium (see Figure 3b), households entering the economy choose this occupation at much higher rates than in the steady state (see Figure 4), with the NRC employment share rising 7 p.p., mainly due to inflows of workers who would otherwise have joined NRM and RM occupations. This higher relative supply of NRC labor is not sufficient to offset the effect of technology on the wages of other occupations. These drop for every occupation in this counterfactual except for the NRC, which experience a 35 p.p. increase in their wage premium relative to the RM. Furthermore, the RM occupation, which loses 9 p.p. of its employment share in the counterfactual relative to 1980, is the occupation with the lowest earnings variance. In contrast, both RC and NRC, the target occupations in this counterfactual, have a much higher level of earnings variance. Therefore total earnings variance rises further via this mechanic.

The drop in the relative price of equipment investment, ξ, is next in terms of importance. Alone, it accounts for 22% of the change in earnings inequality. As the price of equipment falls by 60% over the course of 35 periods, the return on savings rises and spurs capital accumulation. Because of how the production function is specified and of the elasticity estimates we obtain from the data, the marginal productivity of some occupations, like NRC, rises compared to others due to capital-occupation complementarity. As a result, the wage rates of non-routine occupations rise in response to capital accumulation, while those of routine occupations fall. This causes new workers to enter the NRC and NRM occupations, their shares increase by 3.0 and 1.7 p.p., respectively, and they experience a net increase of 13% in wages. In contrast, the RC and RM shares drop by 1.6 and 3.1 p.p., respectively, and their wage rates decrease by 5%. As there is a net inflow into occupations with higher within-group earnings inequality, the final projection also increases due to this mechanic.



(a) NRC employment share (%).

(b) NRM employment share (%).

(c) RC employment share (%).

(d) RM employment share (%).

Figure 4: Decomposition of 2015 Model Prediction of Employment Shares by Occupation.

Note that not all of these forces push toward an increase in earnings inequality. As the wage premium of NRM jobs increases, it lowers inequality. The NRM wage premium rises from 0.62 to 0.68, despite an increase in NRM labor supply, which means that wages between NRM and routine occupations become more compressed due to investment-specific technological change. However, this effect is more than offset by the other forces: The inflow of RC and RM workers moves the mass of workers from the center of the distribution to the bottom and the top, thereby increasing earnings dispersion.

Together, ISTC and LAT account for 65% of the increase in pre-tax earnings dispersion, slightly below the 67% in the baseline model prediction. The most significant forces driving it are the increase in the NRC wage premium, which rises by 50 p.p. to 1.83 of the RM wage rate, and the soaring NRC employment share, which has a higher within-occupation earnings inequality relative to RM occupations.

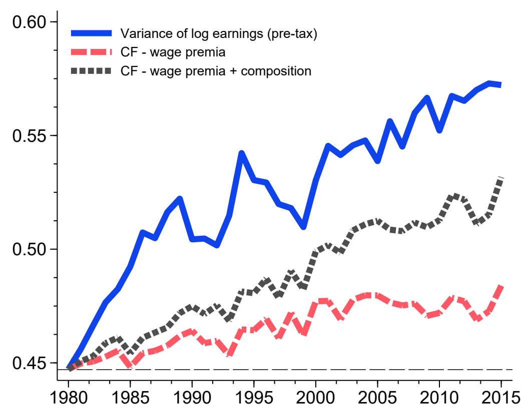
In contrast, other sources of variation in the variance of pre-tax log earnings are much less relevant. Total factor productivity raises capital accumulation and, via capital-occupation complementarity, the wage premium of non-routine occupations and their employment shares. However, it does so by only a few percentage points, producing a limited quantitative impact. Note that this means that TFP is not Hicks-neutral: By raising savings, it raises the wage premium of occupations benefiting from capital-occupation complementarity, along with their employment shares. In isolation, this mechanism accounts for 1.6% of the rise in the variance of log earnings.

Tax changes and the increase in government debt likewise have very limited effects. The more than a quarter drop in progressivity of the labor income tax schedule generates a rise in log earnings variance of only 1.8% of the total observed change. This is close to the magnitude generated by the expansion of government debt-to-GDP, which crowds out capital and reduces the wage premia of non-routine occupations, increasing the dispersion at the bottom of the earnings distribution slightly. Finally, the reduction in capital income taxation increases the reward from capital accumulation, but the extra capital accumulation is not sufficient to produce significant effects on wage premia.

In summary, we find that technological change, especially in the form of LAT and ISTC, generates an increase in pre-tax earnings inequality which is two-thirds of the one observed in the data. In the next sub section, we use a model-free approach to validate the positive conclusions from our model before we move to the study of optimal taxation in Section 7.

6.2 Validating Model-Predicted Changes in Earnings Inequality

The two main forces that change inequality over time in our model are technology-driven changes in wage premia and in the occupational composition of the work force. To verify whether this mechanism can explain two thirds of the increase in inequality between 1980 and 2015, we implement a purely empirical decomposition of the contribution from these two factors.



Note: Log earnings are the natural logarithm of pre-tax weekly earnings usually received by workers. "CF - wage premia" is a sequence of counterfactual variances calculated using the 1980 earnings data and adjusting earnings by occupation to reflect changing wage premia through time. The wage premia are obtained from a regression controlling for education, sex, race, and years of potential experience. Log earnings are adjusted by adding the change in the log wage premium estimate in each period to individual earnings. "CF - wage premia + composition" is a sequence of variances where the individual sample weights of the 1980 earnings distributions are adjusted to reflect changes in employment shares, together with the change in wage premia. The horizontal black dashed line indicates the level of the variance of log earnings in 1980. Source: CPS and authors' calculations.

Figure 5: Empirical Decomposition of the Rise in Earnings Variance.

In Figure 5 the blue line represents the variance of log earnings in the data. The red line is the variance of log earnings we obtain when we adjust the log earnings in 1980 by adding the change in the log wage premium estimate in each period to individual earnings. We find that in isolation the change in the wage premia explains 29% of the change in the variance in log earnings in the data between 1980 and 2015. Finally the black dashed line is the variance in log earnings when we simultaneously change the wage premia and the share of each occupation.

These two factors combined explain 68% of the increase in earnings inequality between 1980 and 2015 — very close to the model prediction.[38]

7 The Implications of Technology-Driven Inequality for Optimal Taxation

In Section 6 we found that the technological transformation between 1980 and 2015 lead to significantly higher earnings inequality. This section addresses the second main question in our paper: How should optimal tax progressivity react to technology-induced increases in inequality? We answer this question both in steady state and taking into account transition dynamics.

We find that both in the 1980 steady state and during the transition the effect of technological change is to lower optimal tax progressivity, even if earnings inequality increases. The main mechanisms driving this result are the increasing productivity of NRC professions, the positive effect of shifting workers to NRC occupations on the wages of lower-paid occupations, and higher returns to wealth stemming from technology driving up productivity.[39] Lower progressivity leads to significant efficiency gains as more households choose the highly productive NRC occupation. However, when more agents choose the higher-paid occupations, it also has the effect of pushing up the wages in the lowest paid occupations. This dampens the redistributive gains from progressive taxation. Finally, higher returns on savings decrease the insurance value of progressivity. In sum, technological change tilts the tradeoffs between efficiency, redistribution, and insurance in favor of flatter taxes, and occupation choice is of first order importance for this result. To show this we conduct the analysis by comparing results in our benchmark model with and without occupation choice.

7.1 Optimal Taxation in the 1980 Steady State

We begin by studying optimal tax progressivity, θ_1, in a steady state calibrated to resemble the U.S. economy in 1980. As in Wu (2021) and Heathcote et al. (2020) (HSV, henceforth), we assume the economy transitions instantaneously to a new long-run equilibrium following a policy change.

The optimal policy experiment is as follows: Holding the level of government consumption, G, fixed at its benchmark 1980 dollar level, we find the socially optimal labor income tax schedule to finance this expenditure. This approach is standard in the literature on optimal taxation in Aiyagari-type OLG models and avoids making assumptions about the utility of public spending.[40] For a given unanticipated, permanent change in tax progressivity, θ_1, we adjust the parameter governing the tax level, θ_0, to clear the government budget constraint. All other taxes remain at their 1980 level. We apply a utilitarian social welfare function and max-

[38]The small difference with the 67% figure with ISTC + LAT + TFP is due to the fact that the model does not match the data with exact precision.

[39]See Table 15 in Appendix I. See also Jordà et al. (2019) for evidence of higher return rates on wealth in the U.S., and Moll et al. (2022) who also argue that technological change raises the return on wealth.

[40]See, e.g., Erosa and Gervais (2002), Conesa and Krueger (2006), Peterman (2016), Wu (2021).

imize the expected utility of an unborn individual, behind the veil of ignorance. Appendix H.1 formally defines the social welfare function and the social planner's problem. Changes in social welfare are measured as the percentage increase in consumption a household born into the old steady state would need to receive in all states of the world to be indifferent to being born into the steady state associated with a new policy.[41]

To dissect our results, we decompose welfare gains across occupations and welfare channels. For the latter, we follow the method of Flodén (2001)[42], which has been widely used in the optimal taxation literature.[43] In a nutshell, the welfare change from choosing a progressivity level $\theta_1^B > \theta_1$ consists of three components: (i) improved insurance against individual risk (*insurance*); (ii) reduced inequality in average lifetime marginal utilities of consumption, leisure, and bequests (*redistribution*); and (iii) an efficiency cost due to reduced labor supply, savings and weaker incentives to choosing highly productive occupations (*efficiency*).[44]

Capital-occupation complementarity and its impact on occupational choices plays a key role in our results. To isolate this interaction, we compare all experiments to an alternative model without occupational choice, where the equilibrium effects on wages and other variables due to changing occupational shares are not present.[45]

Figure 6a plots the change in social welfare as a function of of tax progressivity (solid blue line) alongside its decomposition into redistribution, insurance, and efficiency components. The optimal progressivity level is 0.15—roughly 25% below the actual 1980 value of 0.19 reported by both Wu (2021) and HSV. The welfare gain from moving to this optimum is modest, amounting to a 0.1% increase in consumption. For comparison, HSV find an optimal progressivity of 0.18 for 1980, with near-zero welfare gains.

This result reflects the dominance of the efficiency channel in the neighborhood of $\theta_1 = 0.19$. In contrast, in the model without occupation choice (Figure 6c), the optimal level of progressivity is much closer to the 1980 value, at $\theta_1 = 0.2$. Although all welfare channels are attenuated in the no occupation choice model, the efficiency gains from lowering progressivity are especially diminished. The mechanisms for the impact of occupation choice are as follows.

First, a reduction in progressivity does not increase capital accumulation as much as in the benchmark model.[46] The reason is that in the latter a reduction in progressivity is accompanied by an exodus to more productive occupations (Figure 7c), which generates more investment and raises output by more via capital-occupation complementarity and a higher return to ability in NRC occupations (Figure 7b). In the model without occupational choice,

[41] This is the same definition as presented in Kindermann and Krueger (2022).

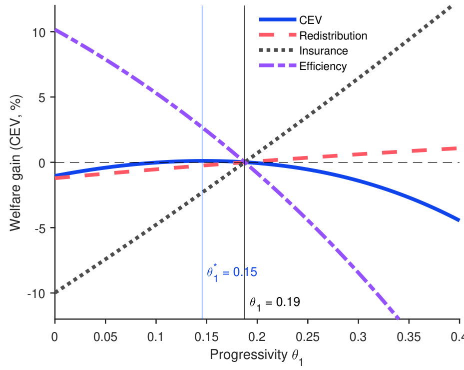
[42] See also Benabou (2002).

[43] See Appendix H.3 for details. An alternative decomposition is proposed in Bhandari et al. (2023), who aim to more directly capture the marginal value of redistribution.

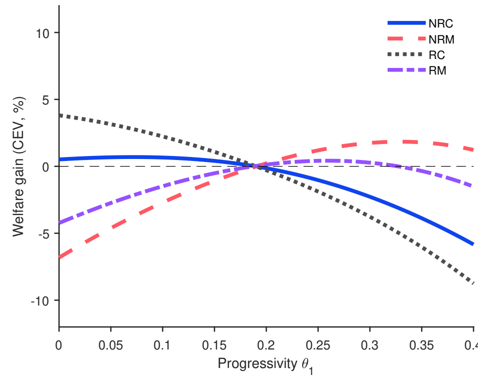
[44] At a given average tax rate, a higher θ_1 implies a higher marginal tax rate, lowering hours worked (see Holter et al., 2019).

[45] This model is identical to the one presented in Section 3 except for two features: (i) agents are randomly assigned to occupations to replicate 1980 employment shares; and (ii) the distribution of ability across occupations is normal, with variance calibrated to match the same moments as the benchmark model in 1980.

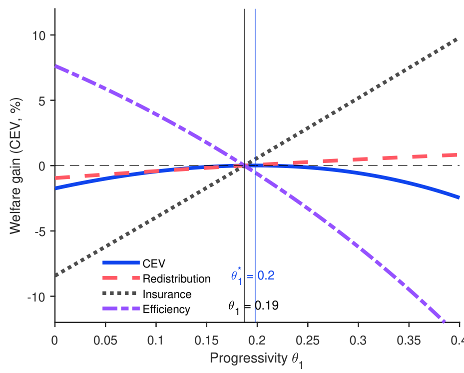
[46] Holter et al. (2019) show that with this tax function, only very low earners will get an increase in their marginal tax rate when progressivity falls. The average tax rate will, however, increase for low-earners.



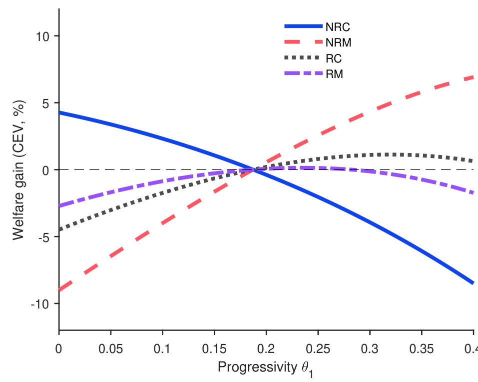
(a) Welfare function and decomposition.



(b) Welfare function by occupation.



(c) Welfare function and decomposition - no occupation choice.



(d) Welfare function by occupation - no occupation choice.

Note: The figure uses the long-run welfare criterion. The top left panel plots social welfare as a function of progressivity, θ_1, under the 1980 calibration, assuming an instantaneous transition to the new steady state. Welfare is measured as the consumption equivalent variation (CEV) for agents entering the economy under the veil of ignorance. The decomposition shows the contributions from redistribution, insurance, and efficiency. Vertical lines mark the current (θ_1) and optimal (θ_1*) progressivity levels. The top right panel shows CEVs by occupation, relative to the 1980 benchmark. The bottom panels report the same results for the model without occupational choice.

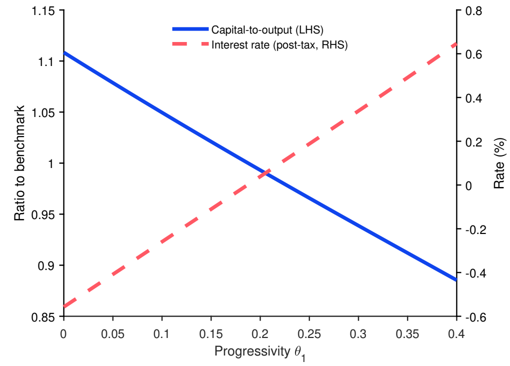
Figure 6: Optimal Tax Progressivity in 1980 for the Long-Run Welfare Measure.

these extensive-margin reallocations are shut down; only hours worked can adjust. As a result, the expansion of the resource constraint is more limited (Figure 8a).

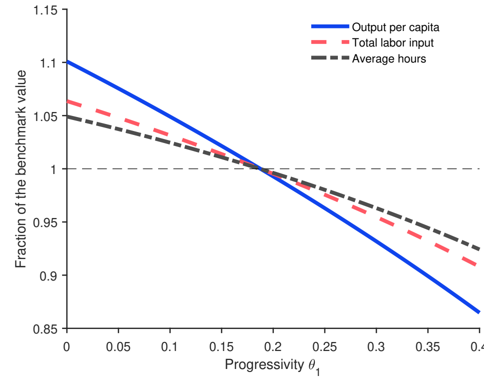
Second, in the model with occupation choice, the reallocation of workers toward the highest-paid occupation (NRC) lowers their pre-tax wage rate (Figure 7d). This mechanism, dubbed the *Stiglitz effect* by HSV, as in Stiglitz (1985), is central to our results. However, we identify an additional dimension to this mechanism: The increase in the share of NRC workers raises the marginal product — and thus the wages — of the other occupations.[47] Note that this result is entirely driven by the data as this is determined by the technical coefficients of the production function estimated in Section 4.

The upshot is that, even though lowering progressivity increases inequality, the redistribution channel remains contained as pre-tax wages of the lower earning occupations rise and counteract this effect. This is a key difference between HSV and our framework: Whereas

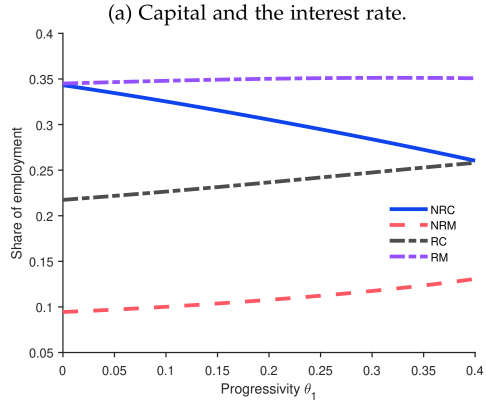
[47]This reallocation also weakens the insurance channel relative to the no occupation choice model due to a composition effect: Routine manual occupations, which contract in membership, have low income risk, while NRC occupations, which expand, are more volatile. This reallocation does not exist in the model without occupation choice.



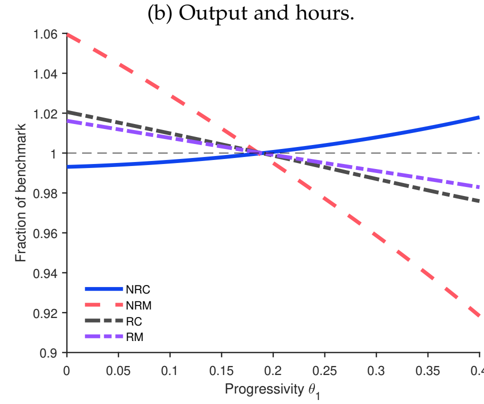
(a) Capital and the interest rate.



(b) Output and hours.



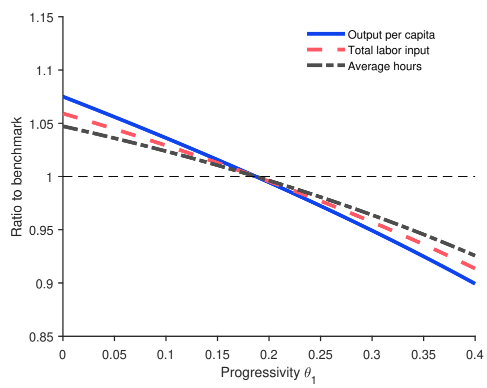
(c) Employment shares by occupation.



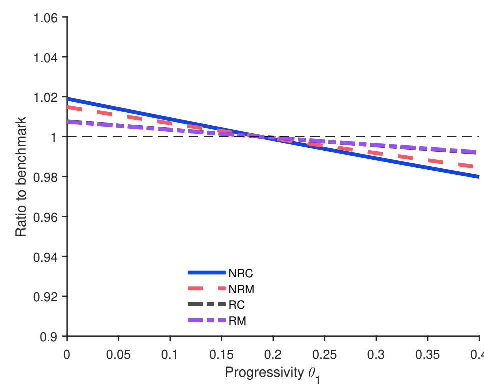
(d) Pre-tax wages by occupation.

Note: Total labor input is the sum of labor efficiency units supplied in the economy. Average hours is the percentage of the labor endowment used to work on average. Wages per occupation are the wage rates per efficiency unit. Each panel shows how prices and quantities in the economy change with respect to progressivity.

Figure 7: Comparative Statics with Respect to Progressivity in 1980.



(a) Output and hours.



(b) Pre-tax wages by occupation.

Figure 8: Comparative Statics with Respect to Progressivity in 1980 - No Occupation Choice.

in their case skill investment is continuous and reversible, our model assumes discrete and irreversible occupation choice. This distinction has two important implications: (i) irreversible occupation choice amplifies the long-run effects of progressivity decisions, and thus using near-term elasticities of effort and labor force participation when deciding on optimal policy (as in Saez, 2001) can ignore important downsides of progressivity in the long-run; and (ii) the extensive nature of occupation choice produces larger aggregate effects than the intensive margin adjustments in HSV. Indeed, in the no occupation choice model, NRC wages rise when progressivity falls, but the wages of other occupations increase much less (Figure 8b).

Figure 6b shows the heterogeneity in optimal progressivity across occupations. NRC workers favor lower progressivity but not a flat tax, as the Stiglitz effect lowers their pre-tax wages, raising their preferred level of redistribution. RC workers benefit from both higher after-tax income and rising pre-tax wages driven by new workers joining NRC occupations rather than RC, making a flat tax optimal for them. In contrast, RM and NRM workers prefer higher progressivity due to the redistributive gains they receive on average.

When the choice of occupation is removed (Figure 6c), the incentives change markedly. Without occupation choice, the Stiglitz effect is weaker, and NRC workers now prefer fully flat taxes, yielding them a welfare gain of nearly 5%. In contrast, the other occupations — except RM — favor substantially higher progressivity, as the benefits from lower progressivity are now greatly diminished.

7.1.1 Accounting for the Transition

Our analysis thus far has abstracted from transition dynamics, assuming the economy moves immediately to a new steady state after a reform. In reality, convergence is gradual, and this matters: Many of the mechanisms discussed above — especially those related to irreversible occupation choice — play out over time. To capture these dynamics, we compute the full transition path following an unanticipated tax reform at $t = 1$, tracking the economy from the initial steady state to its long-run equilibrium. The planner chooses a once and forever value of θ_1 in period t=1 and the tax level parameter, θ_{0t}, is set to clear the budget in every time period. To account for both short- and long-run effects of a reform, we redefine the aggregate welfare criterion as the sum of the expected discounted utility, W_t, of each generation entering the labor market in every period $t \geq 1$, discounted by the policy maker using the discount factor β:

$$\mathcal{W} = \sum_{t=1}^{\infty} \beta^{t-1} W_t. \tag{29}$$

Each generation's welfare is evaluated under the veil of ignorance, prior to the realization of idiosyncratic ability, a_i, the taste shock κ_i, and the starting idiosyncratic risk shock, u_{i1}. Appendix H.2 formally defines the social planner's problem, accounting for the transition.

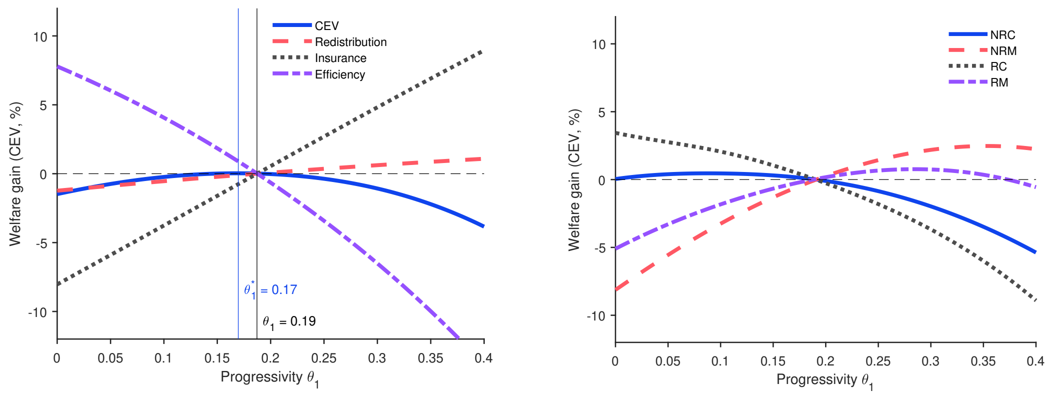
Table 7: Optimal Progressivity in 1980 Accounting for the Transition.

Welfare criterion	Optimal θ_1	CEV (%)
Long-run	0.15	0.11
First generation	0.21	0.05
Aggregate	0.17	0.03

Note: The table shows optimal progressivity in 1980 for the aggregate welfare criterion, for different generations, and the welfare gain from implementing those policies.

Table 7 reports the optimal progressivity and associated welfare gains for the aggregate welfare criterion. The "long-run" criterion corresponds to the welfare of a generation entering the economy in the faraway future, which is equivalent to ignoring the effects of the transition and is therefore equal to the steady state results in the previous sub section. Accounting for short-run effects during the transition raises optimal progressivity from 0.15 to 0.17, still below the actual progressivity of 0.19 and with a lower welfare gain of 0.03% relative to the steady state comparison experiment.

In contrast, the generations born closer to the reform prefer higher levels of progressivity. For example, those born in 1980 would prefer a progressivity of 0.21, instead of 0.17, and their welfare would then rise by 0.05%, measured in consumption equivalents.



(a) Welfare function.

(b) Welfare function by occupation.

Note: The first panel plots social welfare as a function of the progressivity parameter, θ_1, taking into account both short and long-run effects from the transition. Vertical lines mark the 1980 (θ_1) and optimal (θ_1^*) progressivity levels. The decomposition shows the contributions from redistribution, insurance, and efficiency. The second panel panel shows CEVs by occupation, relative to the 1980 benchmark.

Figure 9: Optimal Tax Progressivity in 1980 Accounting for the Transition.

Figure 9a plots the aggregate welfare curve and its decomposition by channel. Relative to the steady-state experiment, the main difference is the weaker efficiency channel: The gains from additional investment and reallocation of households to more productive occupations due to lower progressivity take longer to materialize compared to the case where the economy moves instantaneously to the new steady state. In addition, the insurance channel is weaker as there is a slower shift of employment into occupations with greater risk. However, this latter effect is not enough to offset the flatter efficiency channel, and so optimal progressivity rises compared to the case where short-run effects are ignored.

Figure 9b displays welfare changes as a function of tax progressivity by occupation. Compared to the steady-state analysis (Figure 6b), there is very little change in the preferences of the NRC. They prefere a slightly more progressive tax, possibly because the benefits of lower tax progressivity (capital accumulation) takes time to materialize and will not benefit the first generations as much. For the other occupations it is even more clear that flatter taxes benefit the first generations less. The RC still prefer a flat tax but benefit less from an increase in wage rates following a drop in progressivity. The NRM and RM now benefit more from higher progressivity, given that the general equilibrium effect of worker reallocation on occupation wage rates is weaker early in the transition.

As before, we can compare these results with those of the model with no occupation choice. We find that there are minimal changes with respect to the case where no transitions are considered. Optimal progressivity rises only marginally and the welfare gain is 0.01%. For completeness, we report the results in Appendix I.

However, as previously discussed, the U.S. economy underwent a process of technological change in the decades following 1980 which raised the marginal product of NRC occupations and increased labor income inequality. At first glance, one may intuitively think that this would call for a higher progressivity of the tax system to redistribute the fruits of output growth among lower paid occupations. However, this process also raised the effectiveness of capital accumulation, as equipment investment goods became cheaper. This raises the question we posed at the start of the paper: How do these competing forces affect the optimal progressivity of the tax system? We answer this question in the next sub section.

7.2 The Impact of Technological Change on Optimal Tax Progressivity

In this sub section, we answer the second main question in our paper: How did the technological transformation between 1980 and 2015 affect optimal tax progressivity? To answer this question, we compare the optimal progressivity in 1980 to the optimal progressivity in different counterfactuals, where we replace the 1980 technology parameters with their 2015 counterparts.

We find that the impact of the technological change between 1980 and 2015 is a 40% flatter optimal tax schedule when we account for the transition and a 70% flatter tax schedule in the long-run steady state. It is mainly the change in ISTC that calls for a flatter tax schedule. The results are driven by three main mechanisms which raise the sensitivity of the efficiency channel to progressivity when compared to 1980: (i) lower progressivity generates more investment and output than before, as equipment investment goods become cheaper over time; (ii) workers move to higher paid occupations in greater numbers in response to lower progressivity, as their marginal product rises by more for each additional dollar of saving; and (iii) the stronger reallocation of workers to NRC in response to lower progressivity raises the wages of the other occupations by more. In addition the Stiglitz effect dampens the redistribution channel which would otherwise get steeper with more high earners and increased returns on savings dampens the insurance effect, which would otherwise get steeper when more workers enter the high-risk NRC occupations.

However, this conceals heterogeneity in preferences for progressivity among different generations and occupations. Those entering the labor market in the first years of the transition and poorer workers favor higher levels of progressivity than what is implied by the aggregate welfare criterion.

As in the previous sub section, the policy experiment is a once and for all change in the progressivity, θ_1, in 1980, where the tax level parameter, θ_{0t}, adjusts every period to clear the government budget constraint. When the parameters are varying over time we compute a transition with the relevant benchmark labor income tax system in place (in some of the experiments the benchmark tax parameters are also varying over time) to obtain the level of

Table 8: Impact of Tech Change on Optimal Policy.

	1980	ISTC	LAT	TFP	All Tech
Long-run					
Optimal θ_1	0.15	0.06	0.14	0.15	0.03
CEV (%)	0.11	0.93	0.18	0.11	1.58
Entering in 2000					
Optimal θ_1	0.17	0.13	0.15	0.16	0.11
CEV (%)	0.04	0.21	0.07	0.03	0.44
First generation					
Optimal θ_1	0.21	0.20	0.21	0.21	0.19
CEV (%)	0.05	0.01	0.02	0.04	0.00
Aggregate					
Optimal θ_1	0.17	0.12	0.16	0.17	0.10
CEV (%)	0.03	0.24	0.06	0.02	0.49

Note: The table shows optimal progressivity in the 1980 steady state, and for different technological change scenarios. "ISTC", "LAT", and "TFP" denote scenarios where each source of technological change evolves to its 2015 value in isolation. "All Tech" is a scenario where all sources of technological change simultaneously to their 2015 value.

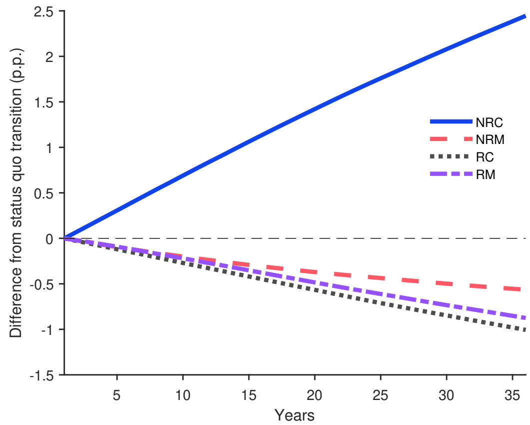
government spending in every period, G_t. Then when we change progressivity we let the tax level parameter, θ_{0t} clear the budget. Table 8 displays the optimal progressivity and the welfare gain from the optimal policy reform for the aggregate welfare criterion (last row), for the 1980 technology parameters and for different technology scenarios. It also reports the breakdown of the aggregate welfare criterion into selected generations, with their preferred progressivity in 1980.

In the "All Tech" scenario, where all types of technology evolve to their 2015 estimates, we find that the optimal progressivity in 1980 is 0.1, with a welfare gain of 0.5%. This is a 40% drop from the 0.17 value presented in Section 7.1 and indicated in the first column. The welfare gain is an order of magnitude larger.

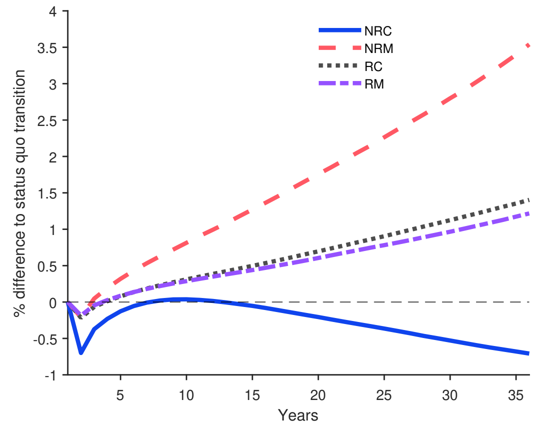
This result is driven by ISTC, which renders equipment investment more efficient (more units of the capital stock per dollar invested) and shrinks optimal progressivity to 0.12. All else equal, ISTC raises the sensitivity of the capital stock to changes in progressivity, as higher paid workers have more disposable income to save. Via capital-occupation complementarity, it also raises the marginal product and after-tax earnings of the NRC by more. As a result, a higher share of new workers entering the labor market will join NRC occupations.

However, because of the Stiglitz effect, these gains are not limited to NRC occupations. As after-tax wages rise in NRC occupations and workers join that category in greater numbers, wages in other occupations rise faster than without ISTC. Therefore, the higher output per capita generated by a larger capital stock and a reallocation of workers to more productive occupations also raises the resources available for consumption by workers at the bottom of the wage distribution, leading to welfare improvements even if the variance of log earnings is higher. Figure 10 shows the impact of implementing optimal progressivity on employment shares and wages relative to the baseline policy. This dampens the negative impact that a reduction in progressivity produces through the redistribution channel.

The effect of ISTC is strengthened by the other two sources of technological change. LAT change puts downward pressure on the wages of all occupations except for the NRC, increasing the flow rate of new workers toward the most productive occupations during the transition. As argued in Section 6.1, TFP growth is not neutral, due to capital-occupation complementarity. Higher TFP generates more saving in response to lower progressivity, which amplifies the effects of ISTC. However, on their own, TFP growth and LAT change produce much weaker effects on optimal progressivity.



(a) Employment shares by occupation.



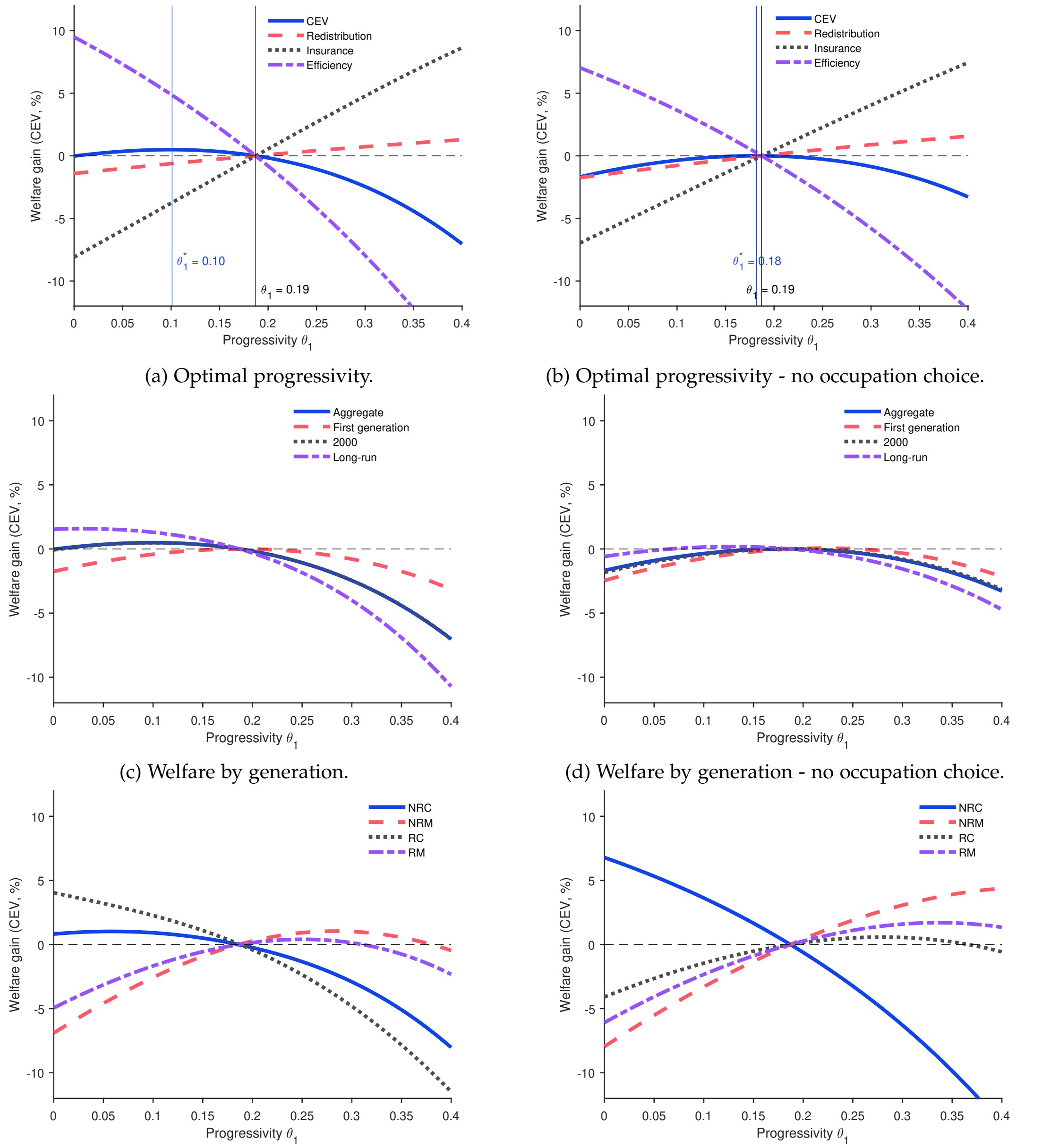
(b) Pre-tax wages by occupation.

Note: The panels show the time series impact of implementing the optimal policy compared to the "All Tech" transition where progressivity is kept at its 1980 level.

Figure 10: Stiglitz Effect in the "All Tech" Counterfactual: Wages at the Bottom of the Wage Distribution Rise as New Workers Join NRC.

These mechanisms contribute to a steeper efficiency channel in response to a change in progressivity with technological change (Figure 11a) compared with the 1980 steady state (Figure 9a). Furthermore, even though reducing progressivity increases after-tax earnings dispersion, the redistribution channel remains nearly unchanged, as the Stiglitz effect counteracts the flow of workers to NRC occupations, keeping pre-tax wage dispersion in check.

The aggregate preference for lower progressivity due to technological change does not imply unanimity. Figures 11c and 11e show the welfare functions of different generations and occupations, respectively, for the "All Tech" scenario. The generation entering the labor market in 1980 prefers a level progressivity equal to the actual one in 1980, as reported in Table 8. This is because the efficiency gains from technological growth and Stiglitz effect unfold only gradually. The first generations will prefer more redistribution compared to the later ones who get the full benefits of technological change. In contrast, generations joining the labor market later, who do not yet know which occupations they will choose, enjoy the benefits of both a higher output per capita and smaller downside risk, as the NRC becomes more productive and the Stiglitz effect kicks in and raises wages at the bottom of the wage distribution. As a result, these generations have a preference for lower progressivity in the "All Tech" scenario: 0.11, with a 0.4% welfare gain, for those entering the labor market in 2000; and 0.03, with a welfare gain of 1.6%, for generations entering the labor market in the long-run.



(a) Optimal progressivity.

(b) Optimal progressivity - no occupation choice.

(c) Welfare by generation.

(d) Welfare by generation - no occupation choice.

(e) Welfare by occupation.

(f) Welfare by occupation - no occupation choice.

Note: θ_1 indicates actual progressivity in 1980. θ_1^* indicates optimal progressivity set once and for all in 1980. The plots labeled “2000” indicate the welfare functions of households entering the labor market in the year 2000.

Figure 11: Optimal Progressivity in the “All Tech” Scenario With and Without Occupation Choice.

Among occupations there is also preference dispersion, as in Section 7.1. Once more, the NRC would like a lower, but still positive progressivity to prevent an excessive flow of new workers into their occupation following a reduction of progressivity (Figure 11e). However, as their marginal product rises much more in response to higher saving than without technology, their preference for progressivity is lower and generates a higher welfare gain. On the other side of the spectrum, RM and NRM wages rise by more from the shift of new workers to other occupations than in the case without technology, so their preference for progressivity is reduced even if it is still much higher than the 0.17 implied by the aggregate welfare criterion.

Finally, we find that occupation choice is a key model ingredient to generate these results. Figure 11b shows the welfare function and its decomposition for the "All Tech" scenario in the model with no occupation choice. In this case, the efficiency channel is much weaker, as the NRC employment share is not allowed to grow in response to higher after-tax earnings. Therefore, the only possible adjustment in response to higher NRC wages is via the intensive margin, which has a much lower impact on output per capita. This leads to a re-balancing of progressivity preferences across occupations (Figure 11f): NRC now prefer flat taxes, which would increase their welfare by 7%. In contrast, all other occupations prefer much higher progressivity, as they are not compensated for lower progressivity by a rise in their pre-tax wage rate due to workers joining the NRC instead of their occupations.

7.3 Optimal Time-Varying Progressivity in the Transition to 2015

In the earlier sections, tax progressivity was fixed at its 1980 level or altered once in the optimal taxation experiments. However, the optimal policy might require a gradual adjustment rather than a permanent change. The increased flexibility of time-varying progressivity will weakly improve welfare relative to the once and for all tax reform. In this subsection we analyze how technological change impacts optimal tax progressivity, allowing the social planner to linearly adjust its starting and ending levels between 1980 and 2015.

Table 9: Impact of Tech Change on Optimal Time-Varying Policy.

	All Tech	Baseline Matched
Entering in 2000		
Optimal θ_1^{start}	0.56	0.64
Optimal θ_1^{end}	0.04	0.10
CEV (%)	1.04	0.48
First generation		
Optimal θ_1^{start}	0.43	0.44
Optimal θ_1^{end}	0.00	0.00
CEV (%)	2.55	1.92
Aggregate		
Optimal θ_1^{start}	0.42	0.42
Optimal θ_1^{end}	0.05	0.12
CEV (%)	1.41	0.49

Note: "Baseline Matched" is a scenario where the baseline transition to 2015 is coupled with changes in the occupation-specific error variances of the idiosyncratic productivity shock such that the within-occupation variance of log earnings is matched. "All Tech" is a scenario where only the technology variables (ISTC, LAT, and TFP) evolve to their 2015 values.

Table 9 shows the results of this exercise for two scenarios, indicated in the columns. "Baseline Matched" is the baseline transition to 2015, described in Section 6, coupled with changes in the occupation-specific error variances of the idiosyncratic productivity shock such that the within-occupation variance of log earnings is matched.[48] This adjustment, which makes up for mechanisms not present in the model that raise total earnings inequality, ensures

[48]Without further adjustment, the total variance of log earnings in 2015 for this scenario is 0.58 compared to 0.57 in the data (a deviation of less than 2%).

the optimal policy exercise reflects the full increase in earnings inequality. "All Tech" is a scenario where only the technology variables (ISTC, LAT, and TFP) evolve to their 2015 values, as discussed in the previous sub section.

We find that the optimal time-varying policy, indicated in the last section of the table, is downward-sloping, starting from 0.42 in 1980 and ending in 0.12 in 2015. If we consider the technological change alone, the optimal policy is still downward sloping, starting at 0.42 but with a lower end point at 0.05.

There are three reasons why optimal progressivity has a downward-sloping profile in both scenarios. First, uncertainty and inequality have a high weight: Households begin their lives at the foot of the age profile of wages, they face a whole lifetime of income uncertainty, and they don't yet know their characteristics or their occupation, given that welfare is evaluated behind the veil of ignorance. Second, gains from higher output per capita via capital-occupation complementarity take time to materialize, as investment prices fall gradually and new generations of households join more productive occupations. Third, because the age profile of wages is upward sloping, households benefit from progressivity lowering through the course of their lives.

Notice this tension in the difference between generations in Table 9. Those entering the labor market in 2000 wish a higher starting point for progressivity than the first generation, which enters in 1980. Considering that we are constraining the optimal policy space to linear policies between 1980 and 2015, the later generations wish for a higher starting point so that when they enter the labor market progressivity is still at a high enough level so that they can enjoy its benefits at the start of their lives. It is still downward-sloping though, so they can keep more of their after tax income later in life, and enjoy an increase in pre-tax wages and saving return rates.

Comparing the optimal policy for the aggregate criterion in the two scenarios we observe that, while the starting points are unchanged, the end point is 50% lower in the "All Tech" scenario. For the generations born later, both the start and ending points are lower in the "All Tech" scenario. As we have learned in the previous sub sections, the effect of technological change on the efficiency channel, which grows stronger over time, lowers progressivity.. However, recall that in the "Baseline Matched" scenario there are other forces at work in addition to technology. We match the increase in the variance of earnings dispersion almost exactly in 2015 using the variances of the innovations to idiosyncratic risk. This implies that idiosyncratic risk rises from 1980 to 2015, which increases the insurance value of progressivity. In addition, government debt-to-GDP trebles in this period, crowding out private capital and weakening the efficiency channel. For these reasons, the reduction in progressivity at the end point is lower in the transition where we take all these extra factors into account.

Comparing the results for the "All Tech" scenario with the analogous in the previous section, where progressivity is set once-and-for-all, optimal progressivity in the end-point is about half of the once-and-for-all value (0.05 versus 0.10). Both are substantially lower than in a scenario with no technical change (0.17 from Table 8). This illustrates how technological change leads to lower progressivity.

Finally, Table 16 in Appendix I presents the same experiments using the model without occupation choice. In this case, the decline in optimal progressivity over time is markedly flatter: the initial values, θ_1^{start}, increase slightly to 0.47 and 0.43 in the "All Tech" and "Baseline Matched" scenarios, respectively, while the terminal values, θ_1^{end}, rise more sharply to 0.09 and 0.18. Once again, this highlights the role of occupation choice in strengthening the efficiency channel, even if those effects are spread out over time.

8 Conclusion

We develop a life-cycle, overlapping generations model with uninsurable idiosyncratic earnings risk, three sources of technological change, a detailed tax system, and occupational choice. Furthermore we have estimated an aggregate production function with capital-occupation complementarity and four types of labor inputs that differ with respect to cognitive complexity and routine task intensity. We use it to calibrate the model to resemble the the U.S. economy in 1980. Feeding in the estimated paths of changes in the price of equipment goods (ISTC), latent occupation-biased technological change (LAT) and TFP growth, we show that technological transformation accounts for two-thirds of the increase in earnings inequality between 1980 and 2015. The main drivers are changes in LAT and ISTC. The former increases the wages of those at the top of the distribution, and reduces them at the bottom. The latter leads to more capital accumulation and higher relative wages of higher-paid occupations, which benefit the most from complementarity with capital.

In isolation, increasing earnings inequality might strengthen the case for redistributive policies. Yet, the technological changes from 1980 to 2015, which greatly increased earnings inequality, also notably decreased optimal tax progressivity. This fall can be almost solely attributed to ISTC. In our model, in addition to the traditional effects of increasing work hours and savings, lower progressivity leads to an inflow of workers into higher-paid occupations, which are more productive with higher ISTC. This raises output but also the wages of those remaining in the occupations at the bottom of the wage distribution, dampening the redistributive benefits of progressive taxation. Finally, technological growth raises the real return rates on saving, making self-insurance easier and thus weakening the insurance role of progressive taxation.

Our work suggests several promising lines for future research. First, while we may find that it is optimal to reduce the progressivity of the labor income tax system, this does not mean that other redistributive policies are not advisable, such as subsidizing access to education or training to enter better-paid occupations (e.g., see Krueger and Ludwig, 2016 and Stantcheva, 2018). There could be important interactions between these policies, the tax system, occupational choice, and wages. Second, we did not study capital or wealth taxation in this paper. However, the importance of capital-occupation complementarities that we demonstrate could have implication for these taxes. Finally, we do not consider job displacement due to technological change and non-participation in the workforce. How would this affect our welfare analysis? Is a progressive tax system the right tool to counter these phenomena, or are targeted measures more appropriate?

References

Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of Labor Economics*. Elsevier.

Acemoglu, D. and Restrepo, P. (2018). Artificial Intelligence, Automation and Work. Boston University - Department of Economics - Working Papers Series dp-298, Boston University - Department of Economics.

Acemoglu, D. and Restrepo, P. (2022). Tasks, automation, and the rise in u.s. wage inequality. *Econometrica*, 90(5):1973–2016.

Acikgoz, O., Hagedorn, M., Holter, H., and Wang, Y. (2022). The Optimum Quantity of Capital and Debt. Working paper.

Aghion, P., Antonin, C., Bunel, S., and Jaravel, X. (2022). Modern manufacturing capital, labor demand, and product market dynamics: Evidence from france. Unpublished manuscript.

Aghion, P. and Jaravel, X. (2023). The labor market effects of the digital revolution: Evidence and theory. *AEA Papers and Proceedings*, 113:219–223.

Aiyagari, S. R. (1994). Uninsured Idiosyncratic Risk and Aggregate Saving. *The Quarterly Journal of Economics*, 109(3):659–684.

Ales, L., Kurnaz, M., and Sleet, C. (2015). Technical change, wage inequality, and taxes. *American Economic Review*, 105(10):3061–3101.

Alon, T. et al. (2018). Earning more by doing less: Human capital specialization and the college wage premium. *manuscript, University of California at San Diego*.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333.

Bakis, O., Kaymak, B., and Poschke, M. (2015). Transitional dynamics and the optimal progressivity of income redistribution. *Review of Economic Dynamics*, 18(3):679–693.

Barro, R. (2000). Inequality and growth in a panel of countries. *Journal of Economic Growth*, 5(1):5–32.

Benabou, R. (2002). Tax and education policy in a heterogeneous agent economy: What levels of redistribution maximize growth and efficiency? *Econometrica*, 70:481–517.

Bewley, T. F. (2000). Has the Decline in the Price of Investment Increased Wealth Inequality? Unpublished.

Bhandari, A., Evans, D., Golosov, M., and Sargent, T. J. (2023). Inequality, business cycles, and monetary-fiscal policy. Unpublished manuscript.

Boar, C. and Lashkari, D. (2021). Occupational Choice and the Intergenerational Mobility of Welfare. NBER Working Papers 29381, National Bureau of Economic Research, Inc.

Boar, C. and Midrigan, V. (2022). Efficient Redistribution. Working paper.

Brinca, P., Holter, H. A., Krusell, P., and Malafry, L. (2016). Fiscal multipliers in the 21st century. *Journal of Monetary Economics*, 77:53–69.

Brinca, P., Homem Ferreira, M., Franco, F. A., Holter, H. A., and Malafry, L. (2021). Fiscal consolidation programs and income inequality. *International Economic Review*, 62(1):405–460.

Cameron, A. and Trivedi, P. (2005). *Microeconometrics*. Cambridge University Press.

Chakraborty, I., Holter, H. A., Hai, R., and Stepanchuk, S. (2017). The real effects of financial (dis)integration: A multi-country equilibrium analysis of europe. *Journal of Monetary Economics*, (85):28–45.

Conesa, J. C., Kitao, S., and Krueger, D. (2009). Taxing capital? not a bad idea after all! *American Economic Review*, 99(1):25–48.

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

Cortes, G. M., Jaimovich, N., Nekarda, C. J., and Siu, H. E. (2020). The dynamics of disappearing routine jobs: A flows approach. *Labour Economics*, 65:101823.

Delaney, K. J. (2017). Droid duties: The robot that takes your job should pay taxes, says bill gates. *Quartz*.

Dyrda, S. and Pedroni, M. (2021). Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Shocks. *Working Paper*.

Eden, M. and Gaggl, P. (2018). On the welfare implications of automation. *Review of Economic Dynamics*, 29:15 – 43.

Erosa, A. and Gervais, M. (2002). Optimal taxation in life-cycle economies. *Journal of Economic Theory*, 105(2):338–369.

Ferriere, A. and Navarro, G. (2024). The heterogeneous effects of government spending: It's all about taxes. *The Review of Economic Studies*, 92(2):1061–1125.

Ferrière, A., Grübener, P., Navarro, G., and Vardishvili, O. (2023). On the optimal design of transfers and income tax progressivity.

Flodén, M. (2001). The effectiveness of government debt and transfers as insurance. *Journal of Monetary Economics*, 48(1):81–108.

Flood, S., King, M., Rodgers, R., Ruggles, S., and Warren, J. R. (2018). Integrated public use microdata series, current population survey: Version 6.0 [dataset].

Gordon, R. (1990). *The Measurement of Durable Goods Prices*. National Bureau of Economic Research, Inc.

Greenwood, J., Hercowitz, Z., and Krusell, P. (1997). Long-Run Implications of Investment-Specific Technological Change. *American Economic Review*, 87(3):342–362.

Guerreiro, J., Rebelo, S., and Teles, P. (2021). Should robots be taxed? *The Review of Economic Studies*, 89(1):279–311.

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

Heathcote, J., Storesletten, K., and Violante, G. L. (2020). Presidential address 2019: How should tax progressivity respond to rising income inequality? *Journal of the European Economic Association*, 18(6):2715–2754.

Heathcote, J., Storesletten, S., and Violante, G. (2017). Optimal tax progressivity: An analytical framework. *Quarterly Journal of Economics*, 134:1693–1754.

Holter, H. A., Krueger, D., and Stepanchuk, S. (2019). How do tax progressivity and household heterogeneity affect laffer curves? *Quantitative Economics*, 10(4):1317–1356.

Hugget, M. (1993). The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies. *Journal of Economic Dynamics and Control*, 17:953–969.

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American economic review*, pages 161–182.

Jordà, Ò., Knoll, K., Kuvshinov, D., Schularick, M., and Taylor, A. M. (2019). The Rate of Return on Everything, 1870–2015*. *The Quarterly Journal of Economics*, 134(3):1225–1298.

Kaplan, G. and Zoch, P. ((2020)). Markups, Labor Market Inequality and the Nature of Work. Working paper, University of Chicago.

Karabarbounis, L. and Neiman, B. (2014). The global decline of the labor share. *The Quarterly Journal of Economics*, 129(1):61–103.

Kina, , Slavik, C., and Yazici, H. (2024). Redistributive capital taxation revisited. *American Economic Journal: Macroeconomics*, 12(2):182–216.

Kindermann, F. and Krueger, D. (2022). High marginal tax rates on the top 1 percent? lessons from a life-cycle model with idiosyncratic income risk. *American Economic Journal: Macroeconomics*, 14(2):319–366.

Krueger, D. and Ludwig, A. (2016). On the optimal provision of social insurance: Progressive taxation versus education subsidies in general equilibrium. *Journal of Monetary Economics*, 77:72–98.

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

Larrimore, J., Burkhauser, R. V., Feng, S., and Zayatz, L. (2008). Consistent cell means for topcoded incomes in the public use march cps (1976-2007). *Journal of Economic and Social Measurement*, 33(2/3):89–128.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In Zarembka, P., editor, *Frontiers in Econometrics*, chapter 4, pages 105–142. Academic Press: New York.

Mendoza, E., Razin, A., and Tesar, L. (1994). Effective tax rates in macroeconomics: Cross-country estimates of tax rates on factor incomes and consumption. *Journal of Monetary Economics*, 34(3):297–323.

Moll, B., Rachel, L., and Restrepo, P. (2022). Uneven growth: Automation's impact on income and wealth inequality. *Econometrica*, 90(6):2645–2683.

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.

Ohanian, L. E., Violante, G. L., Krusell, P., and Ríos-Rull, J. V. (1997). Simulation-based estimation of a nonlinear latent factor aggregate production function. In Mariano, R. S., Schuermann, T., and Weeks, M., editors, *Simulation-Based Inference in Econometrics: Methods and Applications*. Cambridge University Press, Cambridge.

Peterman, W. (2016). The effect of endogenous human capital accumulation on optimal taxation. *Review of Economic Dynamics*, 21:46–71.

Powell, D. and Shan, H. (2012). Income taxes, compensating differentials, and occupational choice: How taxes distort the wage-amenity decision. *American Economic Journal: Economic Policy*, 4(1):224–47.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2):135–146.

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

Slavik, C. and Yazici, H. (2022). Wage risk and the skill premium. *The Economic Journal*, 132:2207–2230.

Stantcheva, S. (2018). Optimal taxation and human capital policies over the life cycle. *Journal of Political Economy*, 125(6):1931–1990.

Stiglitz, J. E. (1985). Equilibrium wage distributions. *The Economic Journal*, 95(379):595–618.

Tauchen, G. (1986). Finite state markov-chain approximations to univariate and vector autoregressions. *Economic Letters*, 20:177–181.

Vom Lehn, C. (2020). Labor market polarization, the decline of routine work, and technological change: A quantitative analysis. *Journal of Monetary Economics*, 110:62–80.

Wu, C. (2021). More unequal income but less progressive taxation. *Journal of Monetary Economics*, 117(C):949–968.

ONLINE APPENDIX

The Appendix is organized as follows. Section A describes the micro data sets used. Section B describes the construction of production factor, price, and output measures. Section C derives the first order conditions of the firms in our model. Section D describes the procedure to estimate the production function. Section E describes the equilibrium concepts for the steady state and the transition. Section F.1 outlines the procedures for computation of steady states and transitions. Section G discusses model validation exercises and robustness checks. Section H describes the welfare criterion used and how to decompose welfare effect of policy changes into the effect on redistribution, insurance and efficiency. Section I reports additional tables and figures.

A Data Sets

A.1 CPS

Imputation. From survey year 1968 to 1975, hours worked in the previous year are not available. We follow Acemoglu and Autor (2011) and impute these by running a regression of hours worked on the previous year on hours worked in the current year, on an indicator variable for whether the individual worked 35+ hours last year or not, on the current labor force status, on an interaction variable between the two previous variables, and on the sector the individual worked in the previous year for the survey years 1976-1978. We then use the estimated equation to assign hours worked in the previous year to the 1968-1975 observations.

Weeks worked last year are not available for 1968-1975 as well. We compute mean weeks worked last year by race and gender for the years 1976-78 for each bracket and impute those means for the 1968-1975 period.

Top-coding. To obtain accurate estimates of earnings inequality and wage premia, we have to account for the top-coding in the CPS earnings data. We use the variables *INCWAGE*, *INCLONGJ* and *OINCWAGE*, in the taxonomy of Flood et al. (2018). We proceed in two steps: (i) identify top-coded observations; (ii) assuming the underlying distribution is Pareto, we forecast the mean value of top-coded observations by extrapolating a Pareto density fitted to the non-top-coded upper end of the observation distribution. For details on the procedure to approximate the tail of a Pareto distribution see Heathcote et al. (2010).

Top-coding thresholds in the ASEC change across variables and time. Information on top-coding thresholds can be found on the IPUMS website. Prior to the 1996 survey year, there is little documentation available regarding the thresholds, but the effective top-coding thresholds are provided by IPUMS based on Larrimore et al. (2008). From 1996 onward, the Census Bureau began reporting top-coding thresholds for a set of income variables.

In addition, the Census Bureau has changed its top-coding procedure through time: from 1996 until 2011, the values for top-coded observations were replaced with values based on the individual's characteristics (so-called cell/group means). From 2011 onward, the Census Bureau shifted from an average-replacement value system to a rank proximity swapping procedure.

Ideally, we would like to use a consistent procedure for handling top-coding across time. However, since the Census Bureau started publishing top-coding procedures in 1996, they drastically reduced public use censoring thresholds. Heathcote et al. (2010) found that the Pareto-extrapolation procedure does not perform well in this case. Therefore, we only apply this procedure until survey year 1995. Heathcote et al. (2010) use the extrapolation until survey year 1999, but we find that this produces a large jump in earnings inequality in the late 90's which does not seem plausible.

Bottom-trimming. According to Flood et al. (2018), there is no publicly available information on bottom-coding thresholds of income variables in the ASEC. To deal with this shortcoming, a common practice in the literature is to select a bottom threshold on earnings for inclusion in the sample. We use the procedure of Heathcote et al. (2010): the final sample only includes observations where the hourly wage is above the minimum threshold of one half of the federal minimum wage in each year (end-year federal minimum wage data for farm and non-farm workers is retrieved from FRED).

Variable definitions. All variables are computed as explained in Acemoglu and Autor (2011).

Sample selection. The population of interest comprises non-military and non-institutionalized individuals aged 16 to 70, excluding the self-employed and farm sector workers. We build two samples, labeled A and B. Table 10 shows the number of records at each stage of the selection process.

Table 10: CPS Sample Selection (survey years 1968-2017).

	Dropped	Remaining
Initial sample		**4,089,617**
Wage > 0.5 $\times$ min. wage	116,608	3,973,009
Sample A		**3,973,009**
Age 25-64	861,598	3,111,411
Hours worked per week last year > 6	19,308	3,092,103
Sample B		**3,092,103**

The initial sample is a cleaned version of the raw data, which excludes individual records which are either: below the age of 16 in the previous year, not part of the universe, not wage workers, did not work in the previous year, have zero or missing weights, missing age, or have positive earnings but no weeks worked in the previous year, or vice-versa. In 2014, two distinct samples were drawn because of sample redesign. We keep the sample which is consistent with previous surveys.

Sample A excludes all records where the hourly wage is lower than one half of the federal hourly minimum wage. We assume that this sample is representative of the (non-institutionalized) U.S. population.

In order validate the data, we compare a set of sample statistics on wages and hours worked to their aggregate (NIPA) counterpart. This is shown on Figure A.1.

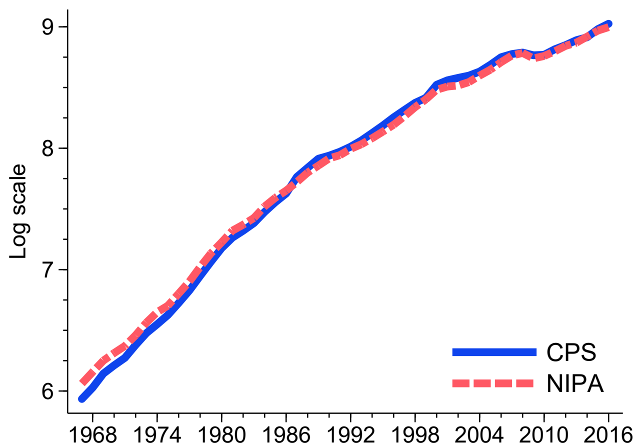
There is an average absolute deviation of 5% between the NIPA (Table 2.1, line 3) and the CPS wage bill. Regarding hours of part and full-time employees, the NIPA series (Tables 6.9B-D, line 2) is lower by 3.3%, on average, and 6.5% after 1986. The BEA uses BLS data to calculate its hours worked series, but the variables are based on the Quarterly Census of Employment and Wages (QCEW) data, rather than on the ASEC variable "usual hours worked per week last year" used in this paper. The total number of full- and part-time employees is much closer to the NIPA series (Table 6.4B-D, line 2), albeit the gap is still 2.7% on average.

Sample B excludes individuals between 25 and 64 years old in the previous year. We consider that 25 years old is a reasonable cutoff age, where individuals' occupation choice has stabilized. According to the BLS, for 2018 the labor force participation rate drops from 65% to 27%, on average, between the 55-64 and the 65 and older age brackets, which justifies our upper bound for inclusion in the sample. We also exclude records where individuals usually worked less than 6 hours per week in the previous year. This is the sample we use to calculate inequality and wage premia statistics. For comparison, Heathcote et al. (2010) have 2,578,035 individual records in their individual-level database, covering the 1967-2005 survey years. This implies that we have around 63,000 records per year, on average, while Heathcote et al. (2010) have 68,000.

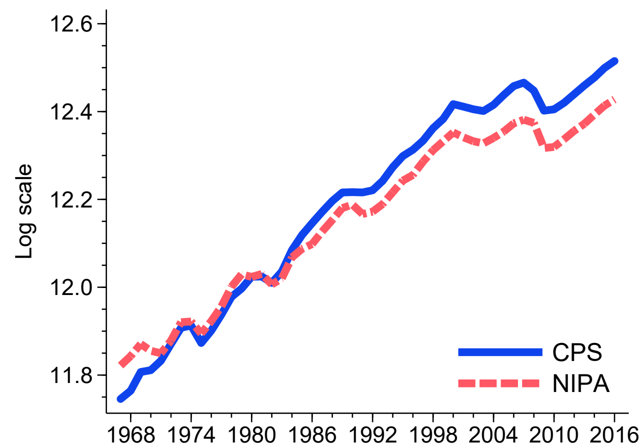
B Measures

B.1 Labor Supply and Wages

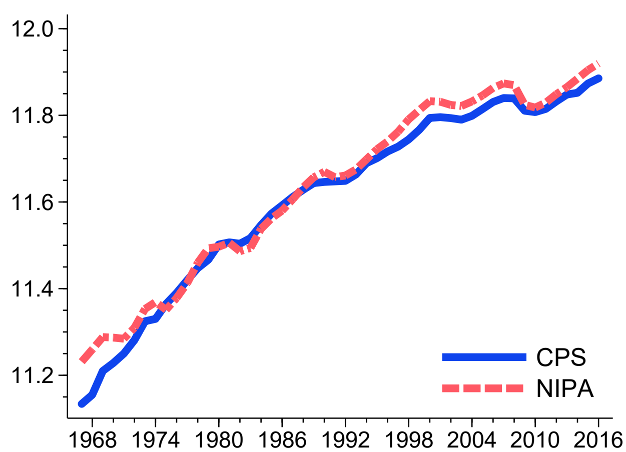
We follow the procedure of Krusell et al. (2000) to build measures of wages and the labor supply for each of the labor categories (NRC, NRM, RC, RM). The sample used for this purpose is the same as the one used for the regression analysis described on section A, apart from the fact that we include workers which did not work full-year or full-time. The reason for this is that in the regression analysis we were aiming to identify the wage premia by observing workers in a similar labor market situation. Here, the aim is to construct measures of labor inputs and wages which will be used in the estimation of the production function. We use these



(a) Wages and salaries



(b) Employee full- and part-time hours



(c) Full and part-time employees

Figure A.1: Comparison between aggregate labor variables in the CPS and in the NIPA.

bins in order to exclude phenomena such as the increased labor force participation of women from the estimation. Since the labor supply of part-time workers contributes to real GDP, it is necessary to account for those. We do not, however, include self-employed individuals in the analysis. In what follows, the subscript t denotes the year and i denotes an individual observation.

For each worker, we record the following variables: hours usually worked per week last year, weeks worked last year, earnings last year, potential experience, race, gender, years of education, occupation category and ASEC weight. Potential experience is divided into 5 five-year groups. Race into white, black and other. There are two sexes. Education is divided into 5 categories: no high school, high school graduate, some college, college graduate, and post college education. Occupation groups are defined as before.

Each worker is assigned to one group defined by the variables described. There are 600 groups, each one denoted by $g \in G$. For each group, we construct a measure of the labor input and labor earnings. The individual labor input is defined as $l_{it} = h_{it} wk_{it}$, where h_{it} is hours usually worked last year and wk_{it} is weeks worked last year. The individual wage is defined as $w_{it} = y_{it}/l_{it}$. Therefore for each group g we define:

$$l_{gt} = \frac{\sum_{i \in g} l_{it} \mu_{it}}{\mu_{gt}},$$

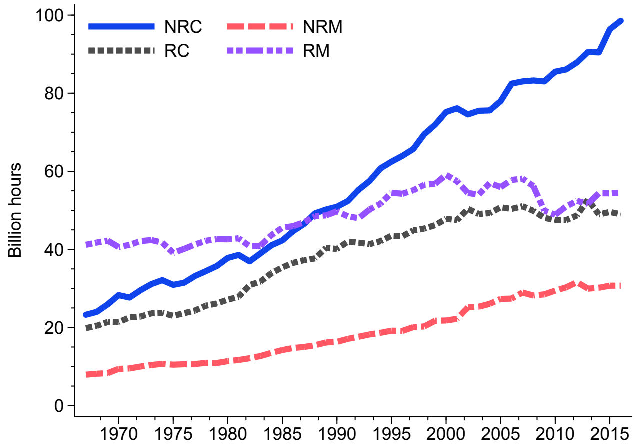
$$w_{gt} = \frac{\sum_{i \in g} w_{it} \mu_{it}}{\mu_{gt}},$$

where μ_{it} is the individual ASEC weight and $\mu_{gt} = \sum_{i \in g} \mu_{it}$. We aggregate the set G of 600 sets into the occupation categories previously defined $o \in \{\mathrm{NRC}, \mathrm{NRM}, \mathrm{RC}, \mathrm{RM}\}$. From this aggregation we obtain total annual labor input per group, $N_{o,t}$, and its hourly wage, w_{ot}. We assume that the groups within a category are perfect substitutes, and for aggregation we use as weights the group wages of 1980. Thus, for each category o, we have:

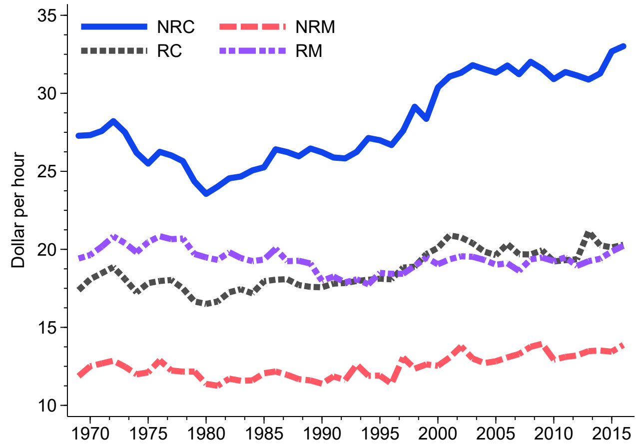
$$N_{ot} = \sum_{g \in s} l_{gt} w_{g80} \mu_{gt},$$

$$w_{ot} = \frac{\sum_{g \in o} w_{gt} l_{gt} \mu_{gt}}{N_{ot}},$$

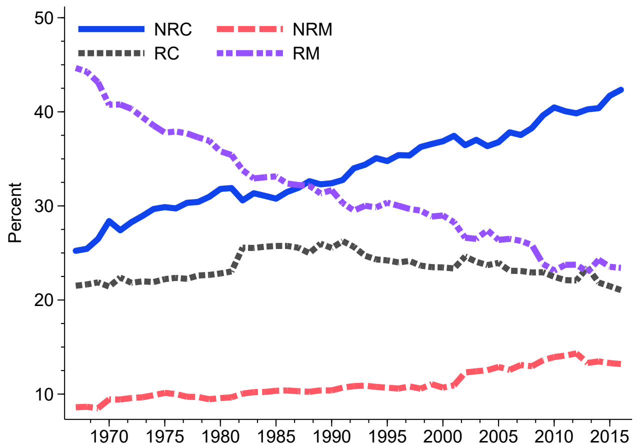
where μ_{it} is the individual ASEC weight and $\mu_{ot} = \sum_{i \in s} \mu_{it}$. This yields a measure of the total labor input in hours by category $(h_{\mathrm{NRC}t}, h_{\mathrm{NRM}t}, h_{\mathrm{RC}t}, h_{\mathrm{RM}t})$, as well as average hourly wages $(w_{\mathrm{NRC}t}, w_{\mathrm{NRM}t}, w_{\mathrm{RC}t}, w_{\mathrm{RM}t})$.



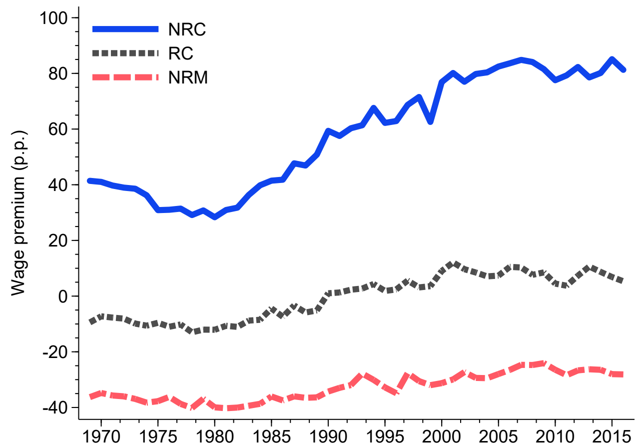
(a) Hours worked.



(b) Hourly Wages at constant Prices.



(c) Share of hours worked.



(d) Wage premia.

Note: Wage premia are obtained as the log difference between the constant composition average wage of each occupation category. Groups for wages are constructed by using a constant composition of individual observable characteristics. The data source is the CPS Annual Social and Economic Supplement. See sections A and B of the Appendix for details.

Figure B.1: Employment and Wages by Occupation Category.

B.2 Capital, Prices and Output

Table 11 shows the definitions of main variables compared with those of Krusell et al. (2000).

Capital. Our main source for capital data are the BEA's fixed asset accounts and the NIPA. We use only private capital in our measure. Nominal investment for each asset category is deflated using the investment price index from the BEA.

Equipment prices. To obtain the price of equipment in each year, we aggregate investment price indices from the BEA fixed asset accounts (Table 5.3.4) across equipment types using a Törqvist index. We then divide the resulting average equipment price by the BLS consumer price index for all urban consumers to obtain the relative price of investment.

Depreciation rates. Obtained using the method by Eden and Gaggl (2018). We use BEA data on the net current cost of the stock of capital, $P_{it}\text{NetStock}_{it}$, and depreciation at current

Table 11: Comparison with Krusell et al. (2000).

Variable	Definition	Definition (KORV)
Output	Business non-farm gross value added	Private domestic product (excluding housing and farm)
Structures	Non-residential structures (private)	Non-residential structures (private)
Equipment	Equipment (private)	Non-military equipment (private)
Equipment price	Equipment price deflator (BEA)	Authors' calculations based on Gordon (1990)

cost, $P_{it}\text{Dep}_{it}$, to compute depreciation rates, which are given by the following formula:

$$\delta_{it} = \frac{P_{it}\text{Dep}_{it}}{P_{it}\text{NetStock}_{it} + P_{it}\text{Dep}_{it}}.$$

We compute average depreciation rates for equipment and non-residential structures, with weights given by the capital stocks at constant prices.

Output. To measure output, we use real gross domestic product in chained 2012 US dollars, retrieved from FRED (FRED code: GDPCA; NIPA code: A191RX).

C First-Order Conditions of the Firms

The first-order conditions of the firms are the following:[49]

$$w_{\text{NRC}t} = \Xi_t \varphi_1 \left[\phi_1 \left(\frac{K_{e,t}}{N_{\text{NRC}t}} \right)^{\frac{\rho_1 - 1}{\rho_1}} + (1-\phi_1) \right]^{\frac{\sigma - \rho_1}{(\rho_1 - 1)\sigma}} [1-\phi_1] \varrho_{\text{NRC}t}, \tag{A-1}$$

$$\begin{aligned} w_{\text{NRM}t} = \Xi_t \varphi_2 & \left[\phi_2 \left(\frac{K_{e,t}}{N_{\text{NRC}t}} \right)^{\frac{\rho_2 - 1}{\rho_2}} + (1-\phi_2) \left(\frac{N_{\text{NRM}t}}{N_{\text{NRC}t}} \right)^{\frac{\rho_2 - 1}{\rho_2}} \right]^{\frac{\sigma - \rho_2}{(\rho_2 - 1)\sigma}} \\ & [1-\phi_2] \left(\frac{N_{\text{NRM}t}}{N_{\text{NRC}t}} \right)^{-\frac{1}{\rho_2}} \varrho_{\text{NRM}t}, \end{aligned} \tag{A-2}$$

$$\begin{aligned} w_{\text{RC}t} = \Xi_t \varphi_3 & \left[\phi_3 \left(\frac{K_{s,t}}{N_{\text{NRC}t}} \right)^{\frac{\rho_3 - 1}{\rho_3}} + (1-\phi_3) \left(\frac{N_{\text{RC},t}}{N_{\text{NRC}t}} \right)^{\frac{\rho_3 - 1}{\rho_3}} \right]^{\frac{\sigma - \rho_3}{(\rho_3 - 1)\sigma}} \\ & [1-\phi_3] \left(\frac{N_{\text{RC},t}}{N_{\text{NRC},t}} \right)^{-\frac{1}{\rho_3}} \varrho_{\text{RC}t}, \end{aligned} \tag{A-3}$$

$$w_{\text{RM}t} = \Xi_t (1 - \varphi_1 - \varphi_2 - \varphi_3) \left(\frac{N_{\text{RM}t}}{N_{\text{NRC}t}} \right)^{-\frac{1}{\sigma}} \varrho_{\text{RM}t}, \tag{A-4}$$

[49] Marginal products are expressed as functions of the ratios between each factor and the non-routine cognitive labor for the purpose of constructing the solution algorithm.

$$r_{s,t} = A_t \alpha \left[\frac{K_{e,t}}{N_{\mathrm{NRC}t}} \right]^{\alpha-1} \Lambda_t^{\frac{\sigma(1-\alpha)}{\sigma-1}}, \tag{A-5}$$

$$\begin{aligned} r_{e,t} = \Xi_t \Bigg[& \varphi_1 \left(\phi_1 \left[\frac{K_{e,t}}{N_{\mathrm{NRC}t}} \right]^{\frac{\rho_1-1}{\rho_1}} + [1-\phi_1] \right)^{\frac{\sigma-\rho_1}{(\rho_1-1)\sigma}} \phi_1 \left(\frac{K_{e,t}}{N_{\mathrm{NRC}t}} \right)^{-\frac{1}{\rho_1}} + \\ & \varphi_2 \left(\phi_2 \left[\frac{K_{e,t}}{N_{\mathrm{NRC}t}} \right]^{\frac{\rho_2-1}{\rho_2}} + [1-\phi_2] \left[\frac{N_{\mathrm{NRM}t}}{N_{\mathrm{NRC}t}} \right]^{\frac{\rho_2-1}{\rho_2}} \right)^{\frac{\sigma-\rho_2}{(\rho_2-1)\sigma}} \phi_2 \left(\frac{K_{e,t}}{N_{\mathrm{NRC}t}} \right)^{-\frac{1}{\rho_2}} + \\ & \varphi_3 \left(\phi_3 \left[\frac{K_{e,t}}{N_{\mathrm{NRC}t}} \right]^{\frac{\rho_3-1}{\rho_3}} + [1-\phi_3] \left[\frac{N_{\mathrm{RC}t}}{N_{\mathrm{NRC}t}} \right]^{\frac{\rho_3-1}{\rho_3}} \right)^{\frac{\sigma-\rho_3}{(\rho_3-1)\sigma}} \phi_3 \left(\frac{K_{e,t}}{N_{\mathrm{NRC}t}} \right)^{-\frac{1}{\rho_3}} \Bigg], \end{aligned} \tag{A-6}$$

where[50]

$$\Xi_t = A_t \left[\frac{K_{s,t}}{N_{\mathrm{NRC}t}} \right]^{\alpha} [1-\alpha] \Lambda_t^{\frac{1-\sigma\alpha}{\sigma-1}}.$$

D Production Function Estimation Method

To estimate the production function, we use the two-step SPML estimator proposed by Ohanian et al. (1997). First, we write the non-linear state space model formally. Next, we briefly describe the methods used to estimate it.

Our non-linear state-space system of equations is of the form:

$$\begin{aligned} &\text{Measurement equations :} && H_t = f(X_t, \psi_t, \omega_t; \theta), \\ &\text{State equations :} && \psi_t = \psi_0 + \psi_1 t + \nu_t. \end{aligned}$$

$f(.)$ contains the labor share equation, the three wage bill equations and the no-arbitrage condition. H_t is thus a (5×1) vector, which is a function of the variables X_t, the log of the unobservable labor quality indices ψ_t, which is a (4×1) vector, and ν_t and ω_t which are (5×1) and (4×1) vectors, respectively, of i.i.d. normally distributed disturbances. Like Krusell et al. (2000), we assume that A_{t+1} and ψ_{t+1} are known when investment decisions are made.

[50] The variable Λ_t is defined as:

$$\begin{aligned} \Lambda_t = {} & \varphi_1 \left(\phi_1 \left[\frac{K_{e,t}}{N_{\mathrm{NRC}t}} \right]^{\frac{\rho_1-1}{\rho_1}} + [1-\phi_1] \right)^{\frac{\rho_1(\sigma-1)}{(\rho_1-1)\sigma}} + \varphi_2 \left(\phi_2 \left[\frac{K_{e,t}}{N_{\mathrm{NRC}t}} \right]^{\frac{\rho_2-1}{\rho_2}} + [1-\phi_2] \left[\frac{N_{\mathrm{NRM}t}}{N_{\mathrm{NRC}t}} \right]^{\frac{\rho_2-1}{\rho_2}} \right)^{\frac{\rho_2(\sigma-1)}{(\rho_2-1)\sigma}} \\ & + \varphi_3 \left(\phi_3 \left[\frac{K_{e,t}}{N_{\mathrm{NRC}t}} \right]^{\frac{\rho_3-1}{\rho_3}} + [1-\phi_3] \left[\frac{N_{\mathrm{RC}t}}{N_{\mathrm{NRC}t}} \right]^{\frac{\rho_3-1}{\rho_3}} \right)^{\frac{\rho_3(\sigma-1)}{(\rho_3-1)\sigma}} + (1-\varphi_1-\varphi_2-\varphi_3) \left(\frac{N_{\mathrm{RM}t}}{N_{\mathrm{NRC}t}} \right)^{\frac{\sigma-1}{\sigma}}. \end{aligned}$$

The model is estimated in two steps: (i) instrument the variables which are potentially endogenous; and (ii) apply the SPML estimator. We assume that the capital stocks, $K_{s,t}$ and $K_{e,t}$, are exogenous at date t. However, we allow for the possibility that date t values of the labor inputs may respond to realization of the technology and labor quality shocks. To instrument these variables, we run a first stage regression of the labor inputs on a constant, current and lagged equipment and structure capital stocks, the lagged relative price of equipment, a trend and the lagged value of the OECD composite leading indicator of business cycles. $\tilde{X}_t$ is the vector of $K_{s,t}$, $K_{e,t}$, the instrumented values of the labor inputs, the depreciation rates and the capital income tax.

The SPML procedure is as follows. Given the distributional assumptions on the error terms, for each t we generate S realizations of the dependent variables, each indexed by i, starting at $t = 1$ in two steps:

$$\begin{aligned} \text{Step 1}: \quad & \psi_t = \psi_0 + \psi_1 t + \nu_t. \\ \text{Step 2}: \quad & H_t^i = f(\tilde{X}_t, \psi_t^i, \omega_t^i, \theta). \end{aligned}$$

In Step 1, we draw a realization of ν_t from its distribution (conditional on our guess of Ω) and use it to construct a date t value for ψ_t. In Step 2, we use our realization of ψ_t, ψ_t^i, together with a draw of ω_t (conditional on our guess of η_ω), to generate a realization of H_t, H_t^i. By using this procedure to generate N realizations, we can obtain first and second simulated moments, respectively, of H_t:

$$\begin{aligned} m_N(\tilde{X}_t;\theta) &= \frac{1}{N}\sum_{i=1}^{N} H_t^i, \\ V_N(\tilde{X}_t;\theta) &= \frac{1}{N-1}\sum_{i=1}^{N}\left(H_t^i - m_N(\tilde{X}_t;\theta)\right)\left(H_t^i - m_N(\tilde{X}_t;\theta)\right)'. \end{aligned}$$

From this procedure, we will obtain $2T$ moments, which we will use to construct an objective function. Denoting the vector of all actual observations of the dependent variables by H^T:

$$L_N(H^T;\theta) = -\frac{1}{2T}\sum_{t=1}^{T}\Big[[H_t - m_N(\tilde{X}_t;\theta)]' V_N(\tilde{X}_t;\theta)^{-1}[H_t - m_N(\tilde{X}_t;\theta)] \ln\det(V_N(\tilde{X}_t;\theta))\Big].$$

The SPML estimator, $\tilde{\theta}_{NT}$, is the maximizer of this expression. It is very important that throughout the maximization procedure of the objective function the same set of $(T \times N)$ random realizations of the dependent variables. Otherwise, the likelihood becomes a random function.

E Equilibrium Concepts

In this Appendix we describe the model solution concepts for the steady state and for the transition path from one steady state to another.

E.1 Definition of a Stationary Recursive Competitive Equilibrium

Letting $\Phi(j, b_{ij}, o_i, a_i, u_{ij})$ be the measure of agents with corresponding characteristics $(j, b_{ij}, o_i, a_i, u_{ij})$, we define a stationary recursive competitive equilibrium as follows:[51]

1. Taking factor prices and initial conditions as given, the value function $V(j, b_{ij}, o_i, a_i, u_{ij})$ and the policy functions, $o(\kappa_o, a_i)$, $c(j, b_{ij}, o_i, a_i, u_{ij})$, $b(j, b_{ij}, o_i, a_i, u_{ij})$, and $h(j, b_{ij}, o_i, a_i, u_{ij})$, solve the working-age household's optimization problem and the occupation choice problem. The value function $V(j, b_{ij}, o_i, a_i)$, and the policy functions, $c(j, b_{ij}, o_i, a_i)$ and $b(j, b_{ij}, o_i, a_i)$, solve the retired household problem.

2. Markets clear:[52]
$$\xi K_e + K_s + B^G = \int b \, d\Phi,$$
$$N_{\text{RM}} = \int n_{\text{RM}} \, d\Phi, \quad N_{\text{RC}} = \int n_{\text{RC}} \, d\Phi,$$
$$N_{\text{NRM}} = \int n_{\text{NRM}} \, d\Phi, \quad N_{\text{NRC}} = \int n_{\text{NRC}} \, d\Phi,$$
$$C + G + \delta_s K_s + \xi \delta_e K_e = F(K_s, K_e, N_{\text{NRC}}, N_{\text{NRM}}, N_{\text{RC}}, N_{\text{RM}}).$$

3. The prices of the production factors equal their marginal products (Equations A-2-A-6 hold).

4. The no-arbitrage conditions (20) and (21) hold.

5. The government budget balances:
$$G + rB = \int \tau_k r(b + \Gamma) + \tau_c c + n\tau_l \left[\frac{hw(j, o, a, u)}{1 + \tilde{\tau}_{ss}}\right] d\Phi.$$

6. The social security system balances:
$$\int_{j>45} \Psi \, d\Phi = \frac{\tilde{\tau}_{ss} + \tau_{ss}}{1 + \tilde{\tau}_{ss}} \left(\int_{j \le 45} hw \, d\Phi \right).$$

[51]The time index is dropped from aggregate variables, given that this is characterization of the steady state.
[52]n_o is the effective labor supply of a household in occupation $o \in O$.

7. The assets of the deceased at the beginning of the period are uniformly distributed among the living:
$$\Gamma \int \omega(j) d\Phi = \int [1-\omega(j)]\, b\, d\Phi.$$

E.2 Definition of a Transition Equilibrium after an Unexpected Shock to the Model Parameters

We define a recursive competitive equilibrium along the transition between steady states as follows. Given the initial capital stocks, $K_{e,0}$ and $K_{s,0}$, the initial distribution of households, Φ_0, and initial government policies, a competitive equilibrium is a sequence of value and policy functions for the household, $\{V_t, o_t, c_t, b_{t+1}, h_t\}_{t=1}^{\infty}$, production plans for the firm, $\{K_{e,t}, K_{e,t}, N_{\text{NRC}t}, N_{\text{NRM}t}, N_{\text{RC}t}, N_{\text{RM}t}\}_{t=1}^{\infty}$, factor prices, $\{r_{e,t}, r_{s,t}, w_{\text{NRC}t}, w_{\text{NRM}t}, w_{\text{RC}t}, w_{\text{RM}t}\}_{t=1}^{\infty}$, government policies $\{\Psi_t, G_t, \theta_{0,t}, \theta_{1,t}, \tau_{c,t}, \tau_{k,t}, B_t\}_{t=1}^{\infty}$, inheritance from the dead, $\{\Gamma_t\}_{t=1}^{\infty}$, and measures $\{\Phi_t\}_{t=1}^{\infty}$, such that for all t:[53]

1. Given factor prices and initial conditions, working age households' optimization problems are solved by value functions $V_t(j, b_{ij}, o_i, a_i, u_{ij})$ and the policy functions, $o_t(j, b_{ij}, o_i, a_i, u_{ij})$, $c_t(j, b_{ij}, o_i, a_i, u_{ij})$, $b_{t+1}(j, b_{ij}, o_i, a_i, u_{ij})$, and $h_t(j, b_{ij}, o_i, a_i, u_{ij})$, and retired households' optimization problems are solved by value functions $V_t(j, b_{ij}, o_i, a_i)$ and policy functions $c_t(j, b_{ij}, o_i, a_i)$ and $b_{t+1}(j, b_{ij}, o_i, a_i)$.

2. Markets clear:
$$\xi_t K_{e,t+1} + K_{s,t+1} + B_t^G = \int b_{t+1}\, d\Phi_t,$$
$$N_{\text{RM}t} = \int n_{\text{RM}t}\, d\Phi_t, \quad N_{\text{RC}t} = \int n_{\text{RC}t}\, d\Phi_t,$$
$$N_{\text{NRM}t} = \int n_{\text{NRM}t}\, d\Phi_t, \quad N_{\text{NRC}t} = \int n_{\text{NRC}t}\, d\Phi_t,$$
$$\int c_t\, d\Phi_t + G_t + K_{s,t+1} + \xi_t K_{e,t+1} = (1-\delta_s)K_{s,t} + \xi_t(1-\delta_e)K_{e,t} + F(.).$$

3. The prices of the production factors equal their marginal products (equations A-2-A-6 hold).

4. The no-arbitrage conditions (20) and (21) hold.

5. The government budget balances:
$$G_t + r_t B_t = \int \left[\tau_k r_t (b_t + \Gamma_t) + \tau_c c_t + h_t w_t \tau_l \left(\frac{h_t w_t(j,o,a,u)}{1+\tilde{\tau}_{ss}}\right)\right] d\Phi_t + (B_{t+1} - B_t).$$

[53] In the quantitative exercises, government final consumption expenditure, G_t, is used to clear the government budget constraint. In the optimal policy exercises, the progressive income tax level, θ_0, is used instead.

6. The social security system balances:

$$\int_{j>45} \Psi_t \, d\Phi_t = \frac{\tilde{\tau}_{ss} + \tau_{ss}}{1 + \tilde{\tau}_{ss}} \int_{j\leq 45} h_t w_t \, d\Phi_t.$$

7. The assets of the dead are uniformly distributed among the living:

$$\int \omega(j)\Gamma_t \, d\Phi_t = \int (1 - \omega(j)) b_t \, d\Phi_t.$$

8. Aggregate law of motion:

$$\Phi_{t+1} = \Delta_t(\Phi_t).$$

F Solution Algorithms

This section describes the algorithms for computing a stationary recursive competitive equilibrium when the economy is in steady state and for computing a sequence of recursive competitive equilibrium in the transitioning economy.

F.1 Stationary Recursive Competitive Equilibrium

To characterize the stationary competitive equilibrium of the model we must find the ratios $\frac{K_s}{N_{NRC}}$, $\frac{K_e}{N_{\text{NRC}}}$, $\frac{N_{\text{NRM}}}{N_{\text{NRC}}}$, $\frac{N_{\text{RC}}}{N_{\text{NRC}}}$, and $\frac{N_{\text{RM}}}{N_{\text{NRC}}}$ which clear capital and labor markets. In addition, we have to fit the tax function, clear the social security budget and find the value of Γ which, given a distribution for the state variable b, uniformly distributes the assets of the dead among the living. G, public consumption of final goods, clears the government budget constraint, except in the case of the optimal taxation exercises where θ_0 is the clearing variable. The algorithm is as follows:

1. Guess $\frac{K_e}{N_{\text{NRC}}}$, $\frac{N_{\text{NRM}}}{N_{\text{NRC}}}$, $\frac{N_{\text{RC}}}{N_{\text{NRC}}}$, and $\frac{N_{\text{RM}}}{N_{\text{NRC}}}$.
2. Obtain the value of $\frac{K_s}{N_{\text{NRC}}}$ which is consistent with the remaining ratios given the no-arbitrage condition (20) using a bisection method. Compute marginal productivities (A-2)-(A-6) with these guesses.
3. Guess ψ_{ss}, Γ, and average earnings.
4. Compute value and policy functions for the retired and active agents by backward induction, given processes for both the transitory, u_{ij}, and permanent shock, a_i. Both shocks are discretized using the Tauchen procedure (Tauchen, 1986), and each has 5 different states. The grids for b_{ij} and h_{ij} have 24 and 100 points, respectively. In between the grid points, the values of the functions are interpolated using cubic splines.

5. Simulate the model for 32,000 agents for each age, where assets holdings are zero for every agent entering the labor market. Occupational choice is simulated by drawing a four-dimensional vector of independent Gumbel shocks for each agent at labor market entry and using the value functions computed in step 4 to solve the maximization problem (1). Obtain total savings (asset demand), $\int b + \Gamma d\Phi$, and quantities of each labor variety supplied, $N_{\text{NRC}}, N_{\text{NRM}}, N_{\text{RC}}, N_{\text{RM}}$.

6. Compute output per capita given implied labor supply of households and using the starting guess of capital stocks. Asset demand must be allocated between government bonds, structure and equipment capital. As government debt is calibrated as a fraction of GDP, we subtract that quantity from assets demand to obtain residual asset demand. We then find the allocation between equipment and structures that satisfies the no-arbitrage condition (20) using a bisection.

7. Obtain implied values for ψ_{ss}, Γ and average earnings. Compare with guesses made in step 4. If the difference between guesses and implied values is within a preset tolerance interval, proceed to step 8. If not, update the guesses of each variable and go back to step 4.[54]

8. Compute the difference between the ratios implied by the labor supply and asset demand of households with the initial guesses. If these differences are within a preset tolerance level, the solution has been reached with sufficient accuracy. If not, update the guesses and go back to step 2.

F.2 Equilibrium on the Transition Path

As is standard when computing transitions, we assume the economy reaches the new steady state after $T < +\infty$ periods. We and set $T = 100$. In our analysis, technological, government policy or other changes are modeled as probability 0 shifts to parameter values from the point of view of households in the economy at $t = 1$, starting from a given steady state (i.e., an MIT shock). Afterwards, the convergence process is deterministic.

To characterize the sequence of competitive equilibria following the shock, we must find the sequence of ratios $\{K_{s,t}/N_{\text{NRC}t}, K_{e,t}/N_{\text{NRC}t}, N_{\text{NRM}t}/N_{\text{NRC}t}, N_{\text{RC}t}/N_{\text{NRC}t}, N_{\text{RM}t}/N_{\text{NRC}t}\}_{t=1}^{T}$, government policies $\{\psi_{ss,t}, G_t, \theta_{0,t}, \theta_{1,t}, \tau_{c,t}, \tau_{k,t}, B_t\}_{t=1}^{T}$, inheritance from the dead, $\{\Gamma_t\}_{t=1}^{T}$, and measures $\{\Phi_t\}_{t=1}^{T}$ such that the conditions described in section E.2 are met. The algorithm is as follows:

[54]Our algorithm uses the homotopy procedure to update all the guesses. That is, if ν is the initial guess and ν' is the value implied by the simulation, then the updated guess is $\nu'' = \nu + a(\nu' - \nu)$, where a is a constant chosen by the researcher which controls the size of the update and the rate of convergence of the algorithm. Due to the complexity of our model, we select homotopy parameters between 0.2 and 0.3 to avoid large steps and, potentially, divergence.

1. Guess sequence of $\{K_{e,t}/N_{\text{NRC}t}, N_{\text{NRM}t}/N_{\text{NRC}t}, N_{\text{RC}t}/N_{\text{NRC}t}, N_{\text{RM}/t}N_{\text{NRC}t}\}_{t=1}^{T}$.

2. Obtain a sequence of $\{K_{s,t}/N_{\text{NRC}t}\}_{t=1}^{T}$ that satisfies the no-arbitrage condition (20) in every period using the bisection method. Compute marginal productivities (A-2)-(A-6) with these guesses.

3. Guess $\{\psi_{ss,t}\}_{t=1}^{T}$, $\{\Gamma_t\}_{t=1}^{T}$, and a sequence of average earnings.

4. Starting from the last period in the transition, compute value and policy functions for retired and active agents by backward induction for every period. Shocks are discretized in the same way as in section F.1, but they are updated every period if there are changes in occupation returns to ability or the standard deviation of the transitory shock.

5. Simulate the model for 32,000 agents for each age, every period, where assets holdings are zero for every agent entering the labor market. Occupational choice is simulated in the same manner as described in section F.1. Obtain a sequence of total savings (asset demand), $\{\int b_{t+1} d\Phi_t\}_{t=1}^{T}$, and a sequence of quantities of each labor variety supplied, $\{N_{\text{NRC}t}, N_{\text{NRM}t}, N_{\text{RC}t}, N_{\text{RM}t}\}_{t=1}^{T}$.

6. Obtain implied $\{K_{e,t}, K_{s,t}\}_{t=1}^{T}$ using the method described in section F.1 for every period in the simulation.

7. Obtain implied sequences of $\{\psi_{ss,t}\}_{t=1}^{T}$, $\{\Gamma_t\}_{t=1}^{T}$ and average earnings. Compare with guesses made in step 4. If the maximum difference between the sequence of guesses and the sequence of implied values is within a preset tolerance interval, proceed to step 8. If not, update the guesses of each variable and go back to step 4.

8. Compute the difference between the sequence of input ratios implied by the labor supply and asset demand of households with the initial guesses. If these differences are within a preset tolerance level, the solution has been reached with sufficient accuracy. If not, update the guesses and go back to step 2.

G Model Validation and Robustness

G.1 Model Validation: Elasticity of Employment Shares to Tax Progressivity

Because the elasticities of employment shares with respect to tax progressivity are non-targeted in the model, we verify that the implied responses are empirically plausible. To do so, we exploit historical variation in U.S. income tax progressivity—our paper's key policy variable—to estimate the response of occupational employment shares in the data.[55] We use local projection methods to trace out the effect of an increase in progressivity on the employment shares

[55] We obtain a yearly time series of progressivity using data from Ferriere and Navarro (2024).

of each of the four occupations, and compare the resulting empirical impulse responses with those generated by the model in response to an MIT shock to progressivity. This exercise provides a quantitative benchmark for assessing whether model-implied elasticities fall within a reasonable empirical range.

We estimate impulse responses using local projections, relating changes in progressivity to future changes in occupational employment shares. This approach allows for flexible dynamics and avoids imposing a specific functional form. Importantly, we treat progressivity as exogenous with respect to short-run labor market outcomes, as shifts in tax policy are typically driven by political and institutional factors rather than contemporaneous occupational trends.

To estimate the dynamic effect of tax progressivity on occupational employment shares, we implement a local projection framework following Jordà (2005). For each forecast horizon $h = 0, 1, \dots, 10$, we estimate the following specification:

$$y^i_{t+h} - y^i_{t-1} = \alpha^i_h + \beta^i_h \theta_{1,t} + \delta^i_h \theta_{0,t} + \lambda^i_h (\theta_{1,t} \times \theta_{0,t}) + \rho^i_h \theta_{1,t-1} + \phi^i_h \theta_{0,t-1} + \psi^i_h Z_t + \varepsilon^i_{t+h}, \quad \text{(A-7)}$$

where y^i_t denotes the employment share (in percentage points) of occupation i in period t. The variable $\theta_{1,t}$ measures tax progressivity, and $\theta_{0,t}$ captures the average tax burden. The model includes their one-period lags as well as an interaction term to control for the non-linear relationship between the tax level and progressivity. The vector Z_t contains controls for occupational wage premia, specifically `wp_nrc`, `wp_rc`, and `wp_nrm`. This flexible specification allows us to trace the impulse response of employment across occupations without imposing a full parametric structure. Standard errors are computed using the heteroskedasticity- and autocorrelation-consistent estimator of Newey and West (1987).

Table 12: Response of Employment Shares by Occupation to a Change in Progressivity in the Data and in the Theoretical Model

	$h = 0$		$h = 10$	
	Data	Model	Data	Model
Non-routine cognitive	-0.098	-0.098	-1.426	-0.919
	(0.297)		(1.169)	
Non-routine manual	-0.181	0.038	0.984	0.3497
	(0.188)		(0.960)	
Routine cognitive	0.350	0.032	1.482	0.314
	(0.574)		(2.803)	
Routine manual	0.385	0.0273	0.361	0.256
	(0.387)		(1.982)	

Note: Reported values correspond to the response (in percentage points) of each occupational employment share to a permanent increase of 0.1 in the progressivity parameter. We present both the short-run response on impact ($h = 0$) and the long-run response after 10 years ($h = 10$). The empirical impulse response functions are estimated using local projections, with heteroskedasticity- and autocorrelation-robust standard errors computed following the method of newey (standard errors shown in parentheses). The model-based impulse responses are simulated from an MIT shock to progressivity, initiated from the 2015 steady state of the model.

Table 12 presents the estimated response of occupational employment shares to a permanent 0.1 increase in the progressivity parameter, comparing empirical estimates from local projections with simulated responses from the model. Results are reported both on impact ($h = 0$) and after ten years ($h = 10$).

In the empirical data, non-routine cognitive (NRC) employment shows a sizable long-run decline of -1.43 percentage points. Non-routine manual (NRM) occupations exhibit a positive long-run response of 0.98 percentage points. In contrast, the long-run responses of routine cognitive (RC) and routine manual (RM) occupations are not statistically distinguishable from zero, given their large standard errors. In the short run, no occupation exhibits a statistically significant response.

The model replicates the direction of the long-run empirical responses across all occupations. It captures the decline in NRC employment and the upward adjustment in NRM, RC, and RM shares, albeit with more muted magnitudes—particularly in the case of RC. In the short run, the model also predicts small, near-zero responses, which mirrors the empirical finding of no statistically significant immediate effect.

Overall, these findings provide meaningful external validation for the model's occupational labor supply elasticities. While the impulse responses were not targeted in calibration, the model successfully reproduces the empirical ordering of responses in terms of sign and relative strength—particularly the pronounced reaction of NRC—thereby supporting the credibility of the substitution patterns embedded in the model's structure.

G.2 Robustness

G.2.1 Stability of the Estimates of the Elasticities of Substitution from the Production Function Estimation Over Time

We examine the stability of the estimated elasticities of substitution by re-estimating the aggregate production function separately for two equal-length subsamples: 1967–1991 and 1992–2016. This exercise is motivated by the literature suggesting that structural features of the labor market may have evolved over time, particularly due to increased occupational specialization and skill complementarity (Heathcote et al., 2020; Alon et al., 2018).

A key macroeconomic development separating these two periods is the onset of widespread digitalization in the early 1990s. This phase saw the increasing diffusion of general-purpose software, databases, and networked computing, which enabled the automation of repetitive yet cognitively demanding tasks. As emphasized by Acemoglu and Autor (2011), such digital technologies disproportionately substituted for workers engaged in routine cognitive tasks, which involve well-defined procedures and rule-based decision-making. These developments are especially relevant for routine cognitive (RC) occupations, which until then had been largely protected from automation.

Table 13 reports the estimated parameters for both periods. The elasticity of substitution between routine manual (RM) labor and the composite of other labor inputs, denoted by σ, declines from 5.636 in the earlier period to 4.106 in the later period. This suggests that RM labor has become increasingly complementary to other inputs—consistent with a growing reliance on integrated production tasks.

For capital-labor substitution, we find that the elasticity between equipment capital and non-routine cognitive (NRC) labor (ρ_1) decreases from 0.462 to 0.160, and a smaller decline is observed for non-routine manual (NRM) labor (ρ_2), from 2.344 to 1.885—both indicating rising complementarity with capital. In contrast, the elasticity of substitution between capital and RC labor (ρ_3) increases markedly from 1.603 to 4.937. This jump is consistent with the idea that capital—especially in the form of software—became increasingly capable of replicating the structured, repetitive tasks typically performed by RC workers in office and administrative roles.

Overall, these findings support the view that technological change has deepened complementarity across labor types and between capital and certain occupations. The notable exception is routine cognitive labor, where the rise in substitutability reflects broader digital transformations beginning in the 1990s.

Aside from this notable case, the estimated elasticities remain relatively stable across time, suggesting that the core structure of the production function is robust to the period of estimation.

Table 13: Elasticities of Substitution Estimates of the Production Function Across Sub-samples

Parameter	Description	Value		
		Baseline: 1967-2016	1967-1991	1992-2016
σ	EOS RM	5.564	5.636	4.106
ρ_1	EOS NRC	0.497	0.462	0.160
ρ_2	EOS NRM	2.055	2.344	1.885
ρ_3	EOS RC	5.029	1.603	4.937

H Social Welfare

This section outlines the welfare criteria used to determine optimal policy. The first part defines the relevant criterion when only long-run considerations are taken into account; specifically, we consider the expected welfare of an unborn household, behind the veil of ignorance. This approach is appropriate for steady-state comparisons. The second part extends the analysis to account for transitional dynamics and all generations born after the start of the transition. For clarity, the subscripts i and j on the state and control variables, and the dependence of value and policy functions on state variables are omitted except where necessary to avoid ambiguity.

H.1 Comparing Welfare in Long-run Steady States

Let $\Phi(j, b, \kappa, a, u)$ be the measure of households with the corresponding characteristics and recall from the main text that the probability of surviving until age j is S_j, which will also be the mass of the generation of age j.[56] The ex-ante lifetime utility of a household before entering the economy is given by:

$$W = \int V(j = 1, b, \kappa, a, u) d\Phi = \mathbb{E}\left[\kappa + \sum_{j=1}^{J} \beta^{j-1}\left[S_j u(c_j, h_j) + (S_j - S_{j+1}) D(b_{j+1})\right]\right], \quad \text{(A-8)}$$

where expectations are taken over the joint distribution of state variables and taste shocks, and κ represents the taste shock incurred upon occupation choice. The objective of the social planner is to solve the following problem:

$$\max_{\{\theta_1, \theta_0\}} W, \quad s.t. \quad \bar{G} = \int \tau_k r(b + \Gamma) + \tau_c c + hw\tau_l\left[\frac{hw}{1 + \tilde{\tau}_{ss}}\right] d\Phi - r\bar{B}. \quad \text{(A-9)}$$

The planner chooses tax function parameters to maximize households' ex-ante lifetime utility while ensuring a balanced budget, with government consumption $\bar{G}$ and debt $\bar{B}$ fixed at their status quo levels.

Let W^A denote the steady state value of (A-8) in the status quo economy and W^B denote

[56]The probability of dying at age j is $S_j - S_{j+1}$.

the corresponding value after a policy reform. Let g denote the uniform percentage change in consumption in all ages and states that equates the expected lifetime utility under the two regimes. Formally, g is the solution to:

$$W^A = W^B(g), \tag{A-10}$$

where $W^B(g)$ is defined as:

$$W^B(g) = \mathbb{E}\left[\kappa + \sum_{j=1}^{J} \beta^{j-1}\left[S_j u(c_j^B(1+g), h_j^B) + (S_j - S_{j+1})D(b_{j+1}^B)\right]\right]. \tag{A-11}$$

The term g thus captures the welfare impact of being born in economy A relative to economy B. A positive g implies a welfare loss, as households entering economy B require compensation to attain the same expected lifetime utility as in the status quo. The optimal policy minimizes g. Substituting (A-11) into (A-10), we derive:

$$\begin{aligned}
W^A &= \mathbb{E}\left[\kappa + \sum_{j=1}^{J} \beta^{j-1}\left[S_j u(c_j^B(1+g), h_j^B) + (S_j - S_{j+1})D(b_{j+1})\right]\right] \\
&= \mathbb{E}\left[\sum_{j=1}^{J} \beta^{j-1} S_j \ln(1+g) + \kappa + \sum_{j=1}^{J} \beta^{j-1}\left[S_j u(c_j^B, h_j^B) + (S_j - S_{j+1})D(b_{j+1}^B)\right]\right] \\
&= \sum_{j=1}^{J} \beta^{j-1} S_j \ln(1+g) + \mathbb{E}\left[\kappa + \sum_{j=1}^{J} \beta^{j-1}\left[S_j u(c_j^B, h_j^B) + (S_j - S_{j+1})D(b_{j+1}^B)\right]\right] \\
&= \ln(1+g) \sum_{j=1}^{J} \beta^{j-1} S_j + W^B \\
g &= \exp\left[\frac{W^A - W^B}{\sum_{j=1}^{J} \beta^{j-1} S_j}\right] - 1,
\end{aligned} \tag{A-12}$$

where the second equality follows from the definition of the utility function, which is separable in consumption and leisure and is logarithmic over consumption.

H.2 Comparing Welfare and Including the Transition

Beyond long-run effects, policy reforms alter welfare for households born during the transition. This section describes an aggregate welfare criterion that accounts for all generations born in $t > 1$.[57] Each generation's welfare is measured as the expected discounted utility, analogous to (A-8), and appropriately discounted depending on when they are born. We

[57] The status quo need not be a steady state; it may itself be a transition under the status quo policy. For simplicity, we describe only the case where the relevant comparison is made with a steady state.

assume the planner maximizes welfare in period, $t = 1$.

Let W_t denote the ex ante utility of a cohort born at time t as defined in (A-8). We define the aggregate welfare of an economy beginning a transition at time, $t = 1$, as:

$$\mathcal{W} = \sum_{t=1}^{\infty} \beta^{t-1} W_t. \tag{A-13}$$

The social planner's problem is therefore:

$$\max_{\{\theta_1,\theta_0\}} \mathcal{W}, \quad s.t. \quad \bar{G}_t = \int \tau_{k,t} r_t (b_t + \Gamma_t) + \tau_{c,t} c_t + h_t w_t \tau_{l,t} \left[\frac{h_t w_t}{1 + \tilde{\tau}_{ss,t}} \right] d\Phi_t - r_t \bar{B}_t, \tag{A-14}$$

where tax rates and policy functions are indexed by t to reflect their time dependence if appropriate. As before, $\bar{G}_t$ denotes the final consumption expenditure of the government in the status quo. If the status quo is itself a transition, $\bar{G}_t$ may also vary over time. Note that τ_l is indexed by t not because $\{\theta_1, \theta_0\}$ change after $t = 1$, but because average earnings, which we use to normalize earnings in the tax function, might.

Denote the status quo by a superscript A and the reformed economy by a superscript B. The consumption equivalent variation g satisfies:

$$\mathcal{W}^A = \mathcal{W}^B(g), \tag{A-15}$$

as in (A-10), but now applied to new aggregate welfare criterion. Following the same steps as before, we derive:

$$\begin{aligned} \mathcal{W}^A &= \sum_{t=1}^{\infty} \beta^{t-1} W_t^B(g) \\ &= \ln(1+g) \sum_{t=1}^{\infty} \beta^{t-1} \sum_{j=1}^{J} \beta^{j-1} S_j + \sum_{t=1}^{\infty} \beta^{t-1} W_t^B \\ g &= \exp\left(\frac{\mathcal{W}^A - \mathcal{W}^B}{\sum_{t=1}^{\infty} \beta^{t-1} \sum_{j=1}^{J} \beta^{j-1} S_j} \right) - 1. \end{aligned} \tag{A-16}$$

As described in section F.2, we approximate the transition path by setting $T = 100$, assuming $W_t^B = W_T^B$ for all $t > T$. Therefore, we approximate the remainder of the infinite series in $\mathcal{W}^B$ by computing $W_T^B/(1-\beta)$ and discounting that quantity to $t = 1$.

H.3 Welfare Change Decomposition à lá Flodén (2001)

One way to break down the changes in welfare into the contributions from changes in *inequality*, *uncertainty* and *efficiency* is the method of Flodén (2001). Define the certainty-equivalent consumption-leisure bundle for a household at age $j = 1$ and taste shock κ as:

$$\sum_{j=1}^{J} \beta^{j-1} u(\bar{c}, \bar{h}) = \kappa + V(1, b, o, a, u) = \tilde{V}, \tag{A-17}$$

where $\bar{c}$ and $\bar{h}$ are constant streams of consumption and labor.[58] Following Flodén (2001), we set $\bar{h}$ to one third, which yields zero utility after retirement. We also remove survival uncertainty on the left-hand side and set the utility of bequests to zero. That leaves only $\bar{c}$ to be determined. Solving (A-17) for $\bar{c}$:

$$\bar{c} = \exp\left(\frac{\tilde{V}}{\sum_{j=1}^{J} \beta^{j-1}} + \chi \frac{\bar{h}^{1+\eta}}{1+\eta} \frac{\sum_{j=1}^{Jret} \beta^{j-1}}{\sum_{j=1}^{J} \beta^{j-1}} \right). \tag{A-18}$$

Define the *cost of inequality* as:

$$\sum_{j=1}^{J} \beta^{j-1} u\left((1 - \rho_{\text{ine}})\bar{C}, \bar{H}\right) = \int \sum_{j=1}^{J} \beta^{j-1} u(\bar{c}, \bar{h})\, d\Phi = W, \tag{A-19}$$

where $\bar{C}$ and $\bar{H}$ are the average of consumption and labor certainty-equivalents, $\bar{C} = \int \bar{c}\, d\Phi$, $\bar{H} = \int \bar{h}\, d\Phi$, and W is as defined in (A-8). Isolating ρ_{ine}:

$$\rho_{\text{ine}} = 1 - \exp\left(\frac{W - u(\bar{H}) \sum_{j=1}^{Jret} \beta^{j-1}}{\sum_{j=1}^{J} \beta^{j-1}} - u(\bar{C}) \right). \tag{A-20}$$

The *cost of uncertainty* is defined as:

$$\sum_{j=1}^{J} \beta^{j-1} u((1 - \rho_{\text{unc}})C, H) = \sum_{j=1}^{J} \beta^{j-1} u(\bar{C}, \bar{H}) = W^{CE}, \tag{A-21}$$

where C is average consumption per capita and H is average hours worked per worker in the economy. Solving for ρ_{unc}:

$$\rho_{\text{unc}} = 1 - \exp\left(\frac{W^{CE} - u(H) \sum_{j=1}^{Jret} \beta^{j-1}}{\sum_{j=1}^{J} \beta^{j-1}} - u(C) \right). \tag{A-22}$$

Denote the variables in the reformed economy with the superscript B. Any policy reform can change the equilibrium levels of both consumption and labor. To measure the welfare effects in terms of consumption only, we define the *leisure-compensated consumption* denoted

[58] For simplicity, we limit the exposition to the case of ex-ante welfare.

by $\tilde{C}^B$:

$$\sum_{j=1}^{J} \beta^{j-1} u(\tilde{C}^B, H^A) = \sum_{j=1}^{J} \beta^{j-1} u(C^B, H^B) = W^{LC}. \tag{A-23}$$

Solving for $\tilde{C}^B$:

$$\tilde{C}^B = \exp\left(\frac{W^{LC}}{\sum_{j=1}^{J} \beta^{j-1}} + \chi \frac{{H^A}^{1+\eta}}{1+\eta} \frac{\sum_{j=1}^{Jret} \beta^{j-1}}{\sum_{j=1}^{J} \beta^{j-1}} \right). \tag{A-24}$$

We now have all the ingredients necessary to define the three separate welfare effects of a change in policy. Let g_{eff} denote the welfare gain in consumption-equivalents from a change in the aggregate levels of consumption and leisure as a result of the policy shift:

$$g_{\text{eff}} = \frac{\tilde{C}^B}{C^A} - 1. \tag{A-25}$$

Denote g_{ine} as the welfare gain from reduced inequality:

$$g_{\text{ine}} = \frac{1 - \rho^B_{\text{ine}}}{1 - \rho^A_{\text{ine}}} - 1. \tag{A-26}$$

Denote g_{unc} as the welfare gain from reduced uncertainty:

$$g_{\text{unc}} = \frac{1 - \rho^B_{\text{unc}}}{1 - \rho^A_{\text{unc}}} - 1. \tag{A-27}$$

Flodén (2001) establishes the following result, which we use to decompose welfare gains into three components:

$$g = (1 + g_{\text{eff}})(1 + g_{\text{ine}})(1 + g_{\text{unc}}) - 1. \tag{A-28}$$

H.3.1 Welfare Change Decomposition Accounting for the Transition

The welfare decomposition for the welfare criterion that accounts for the transition, defined in (A-13), is adapted in the following way. First, the *cost of inequality* is defined as:

$$\begin{aligned} \sum_{t=1}^{\infty} \beta^{t-1} \sum_{j=1}^{J} \beta^{j-1} u((1-\rho_{\text{ine}})\bar{C}, \bar{H}) &= \sum_{t=1}^{\infty} \beta^{t-1} \int_{j=1} \sum_{j=1}^{J} \beta^{j-1} u(\bar{c}, \bar{h})\, d\Phi_t, \\ u((1-\rho_{\text{ine}})\bar{C}) \left[\frac{1}{1-\beta} \sum_{s=1}^{J} \beta^{s-1} \right] &= \mathcal{W} - u(\bar{H}) \left[\frac{1}{1-\beta} \sum_{j=1}^{Jret} \beta^{j-1} \right], \end{aligned}$$

$$\ln(1-\rho_{\text{ine}}) = \frac{\mathcal{W} - u(\bar{H})\left[1/(1-\beta)\sum_{j=1}^{Jret}\beta^{j-1}\right]}{1/(1-\beta)\left[\sum_{j=1}^{J}\beta^{j-1}\right]} - \ln\bar{C}$$

$$\rho_{\text{ine}} = 1-\exp\left[\frac{\mathcal{W} - u(\bar{H})\left[1/(1-\beta)\sum_{j=1}^{Jret}\beta^{j-1}\right]}{\left[1/(1-\beta)\sum_{j=1}^{J}\beta^{j-1}\right]} - \ln\bar{C}\right]$$

For the cost of uncertainty, we have a similar expression:

$$\rho_{\text{unc}} = 1-\exp\left[\frac{W^{CE} - u(H)\left[1/(1-\beta)\sum_{j=1}^{Jret}\beta^{j-1}\right]}{\left[1/(1-\beta)\sum_{j=1}^{J}\beta^{j-1}\right]} - \ln C\right]$$

The right-hand side of the cost of inequality is the weighted average of the expected lifetime utilities for all generations, but expressed in average utility from certainty-equivalent consumption and labor supply. $\bar{C}$ is redefined as:

$$\bar{C} = \frac{\sum_{t=1}^{\infty}\beta^{t-1}\sum_{j=1}^{J}\beta^{j-1}\int_{j=1}\bar{c}\,d\Phi_t}{1/(1-\beta)\sum_{j=1}^{J}\beta^{j-1}},$$

which is the weighted average of certainty-equivalents across time. Φ_t is the distribution of individuals over ability and the idiosyncratic taste shock. Aggregate C and H are defined as the weighted average of annual consumption per capita and annual hours worked per worker across all the periods of the transition, and $\bar{H} = \bar{h}$. All other definitions are as before.

I Additional Tables and Figures

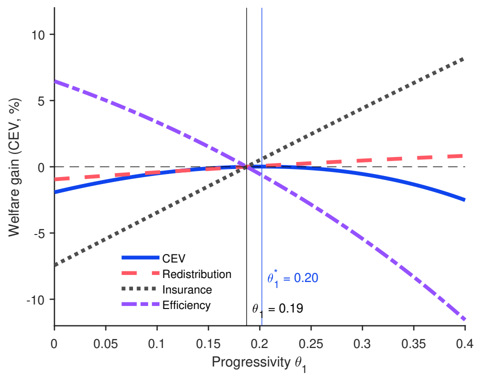
In this Appendix, we report additional tables and figures for the model to complement the results from Section 7.

Table 14 shows the optimal progressivity in 1980, accounting for the short-run effects of the transition. Figure I.1 reports the optimal policy in 1980 for the model with no occupation choice. Table 15 shows aggregate variables in the technological transition scenarios.

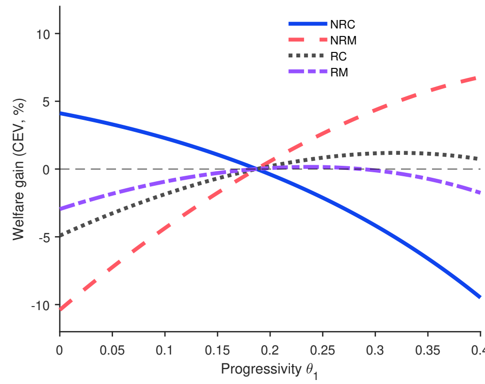
Table 14: Optimal Progressivity in 1980 Accounting for the Transition.

Welfare criterion	Optimal θ_1	CEV (%)
Long-run	0.20	0.01
First generation	0.20	0.01
Aggregate	0.20	0.01

Note: The table shows optimal progressivity in 1980 for the aggregate welfare criterion, for different generations, and the welfare gain in consumption equivalent variation from implementing those policies.



(a) Optimal progressivity.



(b) Welfare by occupation.

Note: The first panel plots social welfare as a function of the progressivity parameter, θ_1, taking into account both short and long-run effects from the transition. Vertical lines mark the current (θ_1) and optimal (θ_1*) progressivity levels. The decomposition shows the contributions from redistribution, insurance, and efficiency. The second panel panel shows CEVs by occupation, relative to the 1980 benchmark.

Figure I.1: Optimal Progressivity in 1980 Accounting for the Transition - No Occupation Choice Model.

Table 15: Variables in Different Tech Change Scenarios.

	1980	ISTC	LAT	TFP	All Tech
Output per capita	1.00	1.22	0.99	1.11	1.38
Capital stock	1.00	2.53	0.96	1.08	2.73
Interest rate (post-tax, %)	1.71	1.93	1.92	1.88	2.39
Wages					
NRC	1.00	1.13	1.01	1.11	1.29
NRM	1.00	1.13	0.88	1.10	1.13
RC	1.00	1.05	0.97	1.10	1.13
RM	1.00	1.05	0.80	1.10	0.92
Employment shares					
NRC	0.31	0.34	0.38	0.31	0.42
NRM	0.11	0.12	0.10	0.11	0.12
RC	0.24	0.22	0.26	0.23	0.23
RM	0.35	0.32	0.26	0.35	0.23

Note: The table shows quantities and prices after 35 years of the start of the transition in different technological change scenarios where optimal progressivity is implemented. Output per capita, the capital stock, and wages are normalized by their 1980 value.

Table 16: Impact of Tech Change on Optimal Time-Varying Policy - No Occupation Choice Model.

	All Tech	Baseline Matched
Entering in 2000		
Optimal θ_1^{start}	0.70	0.70
Optimal θ_1^{end}	0.08	0.15
CEV (%)	1.18	1.23
First generation		
Optimal θ_1^{start}	0.43	0.42
Optimal θ_1^{end}	0.00	0.00
CEV (%)	2.34	1.75
Aggregate		
Optimal θ_1^{start}	0.47	0.43
Optimal θ_1^{end}	0.09	0.18
CEV (%)	0.87	0.80

Note: "Baseline Matched" is a scenario where the baseline transition to 2015 is coupled with changes in the occupation-specific error variances of the idiosyncratic productivity shock such that the within-occupation variance of log earnings is matched. "All Tech" is a scenario where only the technology variables (ISTC, LAT, and TFP) evolve to their 2015 values.