Technology Adoption and Human Capital Accumulation*

Danial Ali Akbari†

January 2023

Abstract

Routine tasks are increasingly becoming automated why labor market skills can depreciate in usefulness and relevance. I develop a task-based framework which incorporates decisions on human capital investment based on the concepts of the psychometric literature on skill formation. The model predicts that labor immiseration – i.e. full automation of the economy – is inevitable unless learning efficiency is improved through capital taxation. While such a scheme can hinder labor immiseration, job polarization, however, is shown to be perpetual and exacerbating as low-index workers are more adversely affected by automation of routine-tasks. The main mechanism for these results are shown to be differences in skill profiles, cross-productivity of skills and the faster accumulation rate of physical vis-à-vis human capital due to advanced skills being more difficult to master.

Key words: Technology adoption; human capital accumulation; task-based framework; labor immiseration, job market polarization
JEL Codes: E24, I26, J24, O33, O41

*I thank my advisers Thomas Fischer, Emiliano Santoro, Joakim Westerlund for support and helpful insights during the process of writing this paper. Moreover, I am deeply grateful to Timo Trimborn and Tobias Broer for detailed comments on the contents of an earlier version of this paper. I also thank Søren Hove Ravn, Alessandro Martinello, Erik Mohlin, Morten Olsen, Pascual Restrepo, Alexandros Rigos, participants in the Lund University macro-metric and theoretical microeconomics seminars, participants in the Copenhagen University macroeconomics seminars and participants in the 24^{th} Annual Conference on Computing in Economics and Finance. I am grateful for financial support from the Thule Foundation through the Skandia Research Program, Tom Hedelius Foundation, Siamon Foundation and Foundation for Economic Research at Lund University. All errors are my own. Corresponding author: Danial Ali Akbari (d.a.akbari@econ.uio.no).

†Department of Economics, University of Oslo.

1 Introduction

Recent literature have reintroduced technological change as a major contributor to labor-displacement with varying impact across performers of different tasks (Acemoglu and Restrepo, 2018*e*; Autor and Salomons, 2018). Indeed, labor market skills can depreciate in usefulness and relevance very suddenly, for instance through introduction of new technologies (automatization and digitalization) or off-shoring of production (Goos, Manning and Salomons, 2014), a process referred to as *human capital obsolescence*. Additionally, the rise of artificial intelligence has given salience to the fear that automating tasks will readily become easier, making a significant share of technological advancements labor-displacing rather than factor-augmenting. Acemoglu and Restrepo (2018*e*) argue that if the arrival rate of automation technology and new labor-intensive tasks are equal then the labor share of production need not diminish if complementary human capital for the new tasks is available. As new non-routine tasks are increasingly difficult to master, however, it is reasonable to suspect that the required human capital may not be able to keep up with the accumulation of physical capital leading to diminished labor shares of production and income in line with the developments of the three recent decades (Elsby, Hobijn and Şahin, 2013; Piketty, 2014; OECD, 2015; Schwellnus et al., 2018). This decrease has been attributed to lower quality-adjusted prices of investment goods such as information and communications technology (Karabarbounis and Neiman, 2014; Dao et al., 2017). Moreover, this reduction in labor share has occurred despite increasing labor productivity and value-added (Madsen, 2014; OECD, 2015; Graetz and Michaels, 2018; Autor and Salomons, 2018).

Indeed this is the first out-of-two main results of this paper. In the framework developed here, since newer tasks are more difficult to master, physical capital accumulates faster than its human counterpart leading to lower prices for technologies that automate routine tasks. This lower price of capital relative to labor provides the favorable conditions for the adoption of automation technology. Consequently, the labor share diminishes continuously along a path towards an inevitable labor immiseration – i.e. full automation of the economy. In other words, by viewing labor as human-capital-augmented, the pivotal channel in determining the technological path of economy, is not the *arrival rate* of technology – as in the framework of Acemoglu and Restrepo (2018*e*) – but rather the conditions for their *institutional adoption*. Labor immiseration can be avoided if improvements learning efficiency is financed through taxing physical capital. Improved learning efficiency increases the rate of human capital accumulation while taxing physical capital reduces its corresponding accumulation rate, creating a balanced growth path.

Thereby, this task-based model also stresses the pivotal role of human capital accumulation and human-capital promoting institutions in determining the growth rate similar to the work of Galor and Moav (2004) and Galor, Moav and Vollrath (2009). For instance, Galor and Moav (2006) argue that, during the era of industrialization, employees and employers found a common interest in seeing the public being broadly educated. Employees demanded education to improve their material conditions, and employers wanted an educated workforce to expand their production capacity and improve their profit margins. Consequently, Galor and Moav (2006) predict the demise of the class structure in the society. The framework developed here, however, predicts an emerging pattern in the class structure in line with recent developments in labor markets – namely *job market polarization* (cf. e.g. Autor, Katz and Kearney, 2006; Goos and Manning, 2007; Goos, Manning and Salomons, 2009; Autor, 2015; Autor and Dorn, 2013; Foote and Ryan, 2015; Harrigan, Reshef and Toubal, 2016). This persisting and deteriorating polarization in the labor market is the second main result of this study. Galor and Moav (2006) argue that class structure will cease as employers and employees find a common interest in a broadly educated public. In the framework here, however, some workers – those with more stock in non-routine skills – are better at keeping up with the technological frontier due to the cross-productivity among their skills. Hence, job polarization and ensuing class divide persists.

As such, my investigations illustrate that labor immiseration is not a merely conceptual construct in the theories of unbalanced growth, rather an exceedingly feasible development path given our evidence on human capital accumulation which is consistent with diminished labor shares. Moreover, while increased educational efficiency can hinder full labor immiseration, it fails to impede job market polarization and instead exacerbates the divide. In other words, the model predicts that labor immiseration is inevitable unless learning efficiency is improved through capital taxation. Moreover, while such a scheme can hinder labor immiseration, job polarization, however, is shown to be perpetual and exacerbating as low-index workers are more adversely affected by automation of routine-tasks. The contribution of this paper is providing a framework that connects these two major structural developments in the economy – i.e. falling labor shares and job market polarization – as a consequence of faster accumulation rate of physical capital and cross-productivity of human capital.

I build on the concepts of task-based framework in Zeira (1998), Autor, Levy and Murnane (2003) and Acemoglu and Restrepo (2018a,b,c,d,e) and augment it with considerations on agents' human capital investments. It has a learning-by-learning structure – i.e. workers increase their stock of human capital by devoting attention to learning. More precisely, workers divide their attention budget between labor – that will generate instantaneous income – and learning – that will be added to the stock of human capital, increase the augmented wage rate and thus indirectly provide labor gains by generating future earnings. The human capital accumulation scheme is microfounded on the psychometric literature regarding skill formation (cf. e.g. Cunha et al., 2006; Cunha and Heckman, 2007, 2008; Almond and Currie, 2011; Cunha, Heckman and Schennach, 2010; Helmers and Patnam, 2011; Graff Zivin and Neidell, 2013). Consequently, the predictions of this model qualitatively match the empirical literature on returns to education as well. This matching is crucial so that the incentive structure prompting workers to pursue education is realistic in the model. Moreover, agents are aware that risk of skill obsolescence due to automation entails future income streams being tentative. Workers also face uninsurable idiosyncratic wealth shocks. Hence, the workforce is heterogeneous in both human capital endowment (*ex-ante*) and wealth (*ex-post*). Furthermore, I follow Itskhoki and Moll (2019) in providing an intergenerational interpretation of the model which then predicts dynastic human capital as documented by Long and Ferrie (2007, 2013), Lindahl et al. (2014, 2015) and Turner et al. (2018).

In short, this project produces a task-based framework in a heterogeneous distributional setting which unifies several empirical findings in the literature concerning automation and its consequences while being consistent with the empirical evidence on skill formation and returns-to-education. Below, in (I), I list the main components of the psychometric literature on skill formation upon which the learning technology of this model is based. Then in (II) to (IV), I present a list of empirical evidence pieces and stylized facts with which the model developed here is consistent.

I. *Evidence from the Psychometric Literature on Skill Formation.* These are incorporated as assumptions, why they are marked as a_1, a_2 and a_3.

   a_1 Self-productivity of human capital: Skills produced at one stage augment the ones attained at later stages (Cunha et al., 2006; Cunha and Heckman, 2007, 2008; Almond and Currie, 2011; Carneiro, Heckman and Vytlacil, 2011; Graff Zivin and Neidell, 2013).

   a_2 Dynamic complementarity: Skills produced at one stage raise the productivity of human capital investment at subsequent stages (Cunha et al., 2006; Cunha and Heckman, 2007, 2008; Cunha, Heckman and Schennach, 2010; Almond and Currie, 2011; Graff Zivin and Neidell, 2013).

   a_3 (Local) cross-productivity of human capital: Stock in one skill eases acquisition of other (related) skills, and vice-versa (Helmers and Patnam, 2011; Cunha and Heckman, 2007, 2008; Carneiro, Heckman and Vytlacil, 2011).

II. *Empirical Evidence on Returns to Education and Skill Profiles*

1. Diminishing marginal internal rates of return (IRR) to education (Card, 1999; Heckman, Humphries and Veramendi, 2018) and subsequent difference between average and marginal IRR to schooling (Heckman, Lochner and Todd, 2006, 2008; Heckman, Schmierer and Urzua, 2010).
2. Heterogeneity in IRR across time and skill profile (Psacharopoulos and Patrinos, 2004, 2018; Heckman, Lochner and Todd, 2008), and self-selection into schooling based on realized returns (Heckman, Schmierer and Urzua, 2010; Carneiro, Heckman and Vytlacil, 2011).
3. More educated agents are more prone to adopting new complex technology (Comin and Hobijn, 2004; Foster and Rosenzweig, 2010).
4. Dynastic human capital as the result of intergenerational transmission (Long and Ferrie, 2007, 2013; Lindahl et al., 2014, 2015; Turner et al., 2018).

III. *Empirical Evidence on the Job Market*

5. Routine-biased technological change (RBTC) (cf. e.g. Autor, Levy and Murnane, 2003; Goos, Manning and Salomons, 2009, 2014; Acemoglu and Autor, 2011; Adermon and Gustavsson, 2015; Cortes, Jaimovich and Siu, 2017; Hershbein and Kahn, 2018).
6. Job market polarization (cf. e.g. Autor, Katz and Kearney, 2006; Goos and Manning, 2007; Goos, Manning and Salomons, 2009; Autor, 2015; Autor and Dorn, 2013; Foote and Ryan, 2015; Harrigan, Reshef and Toubal, 2016).

IV. Growth and Productivity Puzzles

7. Falling labor share of national income (Elsby, Hobijn and Şahin, 2013; Piketty, 2014; OECD, 2015; Schwellnus et al., 2018) despite increasing labor productivity and value-added (David, 1990; Madsen, 2014; OECD, 2015; Graetz and Michaels, 2018; Autor and Salomons, 2018) and its relation to lower quality-adjusted prices of investment goods such as IT and computers (Karabarbounis and Neiman, 2014; Dao et al., 2017).

In the rest of this paper, I will continuously refer to the empirical observations documented in the introduction list by expressing the corresponding number of the specific piece of evidence in brackets. Hence, [a_1] refers to the first piece of evidence, and so on. The consistency of the model with the empirical evidence on returns to education and skill profiles (II) follows directly from assumptions grounded in the psychometric literature on skill formation (I). This consistency is important so that the incentive structure prompting workers to pursue education is empirically sound in the model. In the intergenerational interpretation of the setting here when discussing dynastic human capital, I adhere to viewing human capital stock also representing both sharpness of skills and extent of networks for the sector employing said skills. RBTC [5] follows from ordering tasks from most to least routine and assuming that labor has comparative advantage in non-routine tasks over machines. This setting borrows directly from Autor, Levy and Murnane (2003) and Acemoglu and Restrepo (2018e). Job market polarization [6] follows mainly from labor's comparative advantage in non-routine tasks and the cross-productivity of skills [a_3]. Falling labor share of national income due to lower price of capital [7] – and the labor immiseration as a worst case – follows from RBTC and faster rate of physical capital accumulation relative to human capital. In other words, physical capital becomes abundant more rapidly compared to human capital driving down its price consistent with the empirical findings listed above [7].

The rest of the paper will have the following structure. The related literature is discussed in Section 2. Then in Section 3 the model environment is elaborated. To provide some intuition, in Section 3.1 I will first present an expository partial-equilibrium bivariate static model of skill diversification under risk for obsolescence. In Section 3.2 I then develop the full dynamic general equilibrium model with a continuum of tasks. Section 3.3 specifies the balanced growth path and illustrates the inevitability of labor immiseration provided increasing difficulty in learning more advanced tasks. Section 3.4 shows that while investments in education (through tax revenues from capital) offers a remedy for labor immiseration, any balanced growth path will perpetuate and exacerbate the ensuing job market polarization. In Section 4 I discuss the assumptions and their role in the predicted dynamics. Finally, in Section 5 I will summarize the key results, make some concluding remarks and put forth some future issues for further research.

2 Related Literature

This study is related to three strains of modern economics: the literature on unbalanced growth, the one on the task-based framework and the macroeconomic literature human capital accumulation. One of the earliest and most widely-cited work on labor immiseration vis-à-vis unbalanced growth within the economic literature is Baumol (1967) which anticipates total obsolescence of sectors undergoing automatization. Several authors have recently built on these insights in order to match – with varying degrees of alarmism – the empirical evidence on RBTC [5], job market polarization [6] and the recent dissipation of middling (i.e. middle-income) tasks (cf. e.g. Acemoglu and Autor, 2011; Autor, 2013, 2019; Autor and Salomons, 2018). For instance Berg, Buffie and Zanna (2018) find that in most conceivable cases automation has strong positive effects on growth and a negative impact on equality – even in the optimistic case of robots only being used for a subset tasks or immunity of certain sectors to such automation technology. A more pessimistic view, Susskind (2017) predicts full immiseration of labor through a process of *task encroachment* – that is, as 'advanced capital' accumulates the subset of tasks performed by labor diminishes over time and approaches zero. Feng and Graetz (2019) find evidence of such encroachment by studying job training requirements. They illustrate that employment have been polarized by initial occupational training requirements and shifted towards more complex occupations; and that the relationship between complexity and employment growth is weakest among occupations with low training requirements.

Acemoglu and Restrepo (2018*e*) is the most comprehensive work on capturing the insights of the task-based framework – the second set of related literature. Acemoglu and Restrepo (2018*e*), and the extension in Acemoglu and Restrepo (2018*c*), build on the theoretical grounds laid by Zeira (1998) and the empirically-backed conceptual framework described in Autor, Levy and Murnane (2003). The task-based approach offers a more radical form of creative destruction than that of the early Schumpeterian growth framework (cf. e.g. Aghion, 2002; Aghion, Akcigit and Howitt, 2014). In the latter approach the destructive part of technological change is limited to less productive intermediate firms being replaced by ones which are more so or better at patenting their innovations. In the task-based framework, on the other hand, two kinds of destruction patterns are present: labor-skill obsolescence (as a result of automation) and automated tasks becoming obsolete (following new and ground-breaking technology), which both are potentially more ruinous than a number of firms going bust. A similar approach is made in Acemoglu and Restrepo (2018*c*) although for the distinguished low- and high-skill automating capital stocks.

Finally, this investigation relates to the field of human capital accumulation which is studied in several recent macroeconomic investigations such as Huggett, Ventura and Yaron (2006, 2011), Wallenius (2011), Ludwig, Schelkle and Vogel (2012), Guvenen, Kuruscu and Ozkan (2014), Krebs, Kuhn and Wright (2015) and Ali Akbari and Fischer (2020). None of these, however, consider investment in a heterogeneous set of human capital, the novel contribution of this current study which employs a *learning-by-learning* human capital accumulation scheme. In learning-by-learning models, agents devote a portion of their time (Becker, 1985, 2009), attention (Becker and Murphy, 1992), human capital stock (Ben-Porath, 1967; Rosen, 1976, 1983) or a combination of the three to learning (see e.g. Heckman, 1976, for human-capital-augmented attention allocation). Subsequently, models of this type emphasize the resulting *foregone earnings* during education acquisition (some empirical studies are Heckman and Robb Jr, 1985; Card, 1999; Ginther, 2000; Chabé-Ferret, 2015). The task-based model developed in this study, thereby stresses the pivotal role of human capital accumulation and human-capital promoting institutions in determining the growth rate similar to the work of Galor and Moav (2004) and Galor, Moav and Vollrath (2009). Galor and Moav (2006) predict the demise of the class structure in the society, as employees and employers find a common interest in seeing the public being broadly educated. The framework developed here, however, some workers – those with more stock in non-routine skills – are better at keeping up with the technological frontier due to the cross-productivity among their skills. Hence, job polarization and ensuing class divide persists.

Conceptually, nevertheless, the most related papers are Autor, Levy and Murnane (2003) and Acemoglu and Restrepo (2018*e*) on the task-based framework. I incorporate their ordering of tasks along their degree of routineness and assume comparative advantage of labor in non-routine tasks. This paper also relates closely to the discussion in Galor and Moav (2006) which emphasizes the role of human-capital producing institutions in alleviating inequality. Moreover, concepts on skill formation are borrowed from Heckman, Lochner and Todd (2006) and Heckman, Humphries and Veramendi (2018). There are a number of recent theoretical papers which investigate the impact of automation and digitalization on growth and factor shares, namely Aghion, Jones and Jones (2019), Prettner and Strulik (2020) and Grossman et al. (2020). The relation of their frameworks and results to mine, however, is best understood once its mechanics are explicated. A detailed discussion is found in Section 4, and parallels are drawn continuously in the text where relevant.

3 Model Environment

In this section I develop a model of skill diversification. First, in Section 3.1 an expository static model is developed where the agent divides her attention between learning two types of tasks and working with each skill's corresponding task. This simplified model illustrates the key role of endowments in the types of skill and their corresponding obsolescence probabilities for which skill profile agent's choose. A key result is that in absence of skill obsolescence agents exclusively specialize in one of the skills.

Next, in Section 3.2 a full dynamic model is developed where agents divide their budget of attention between learning a continuum of skills and executing their corresponding tasks. I order the tasks from most to least routine in line with the framework in Acemoglu and Restrepo (2018*e*). Machines have comparative advantage in routine tasks and labor in non-routine ones. At the same time, non-routine tasks are more difficult to master. The insights of the expository model carry over while also allowing for analyzing the technological path of the economy as time progresses.

Thereafter, in Section 3.3 I outline one of the main results of the paper, namely the inevitability of labor immiseration – i.e. full automationof the economy. Given difficulty in mastering new and non-routine tasks, adoption of new technology over time becomes less profitable while automation takes up an increasing share of production. Hence a key insight is that labor immiseration is inevitable due to faster accumulation physical compared to human capital.

Finally, I show that we can remedy labor immiseration with taxing capital and investing the revenues in increasing learning productivity. However, workers adept at non-routine reap the benefits of these investments disproportionately. They are also safe from skill obsolescence due to cross-productivity of human capital. As such in Section 3.4, I illustrate how such schooling policies lead to perpetuate and intensifying job market polarization.

3.1 An Expository Dichotomous Static Model of Skill Diversification

We begin with an expository model in a static two-skill environment. We assume that the agent wants to maximize expected utility. There are two skills $j \in \{1, 2\}$ generating income streams $y_j = w_j h_j \ell_j$ where w_j is the wage level, h_j is the stock of human capital and ℓ_j is (attention to) labor, all pertaining to skill j. The sum of the income streams equal consumption c. The attention budget of the agent is given by

$$a_1 + \ell_1 + a_2 + \ell_2 = 1 \tag{1}$$

where a_j is attention to learning as stated previously. Human capital is accumulated according to the following learning function:

$$h_j = g_j(a_j)h_{j,0}, j = 1, 2, \qquad g_j(0) = 1, g_j' > 0, g_j'' < 0 \text{ for } a_j \in [0, 1], \tag{2}$$

where $h_{j,0}$ is initial stock of human capital in skill j. The characteristics assumed for the learning function $g_j' > 0$ and $g_j'' < 0$ is standard in the literature (cf. e.g. Willis, 1986; Kalemli-Ozcan, Ryder and Weil, 2000; Cunha et al., 2006).

This modeling choice of learning technology is motivated by the psychometric literature which offers two main empirical insights that are relevant to this study. First, the quality and magnitude of early childhood investments have large consequence for skill acquisition later in life (Almond and Currie, 2011; Graff Zivin and Neidell, 2013). Second, different skills complement each other - higher skill in one indicates higher productivity in acquiring the other (Cunha, Heckman and Schennach, 2010; Helmers and Patnam, 2011). Cunha et al. (2006) and Cunha and Heckman (2007, 2008) provides a framework for expressing these results: through three main concepts. First, *self-productivity* of human capital which means that skills produced at one stage augment the ones attained at later stages. Second, *dynamic complementarity* indicating that skills produced at one stage raise the productivity of human capital investment at subsequent stages. Finally, *cross-productivity* of human capital implying that stock in one skill eases acquisition of other skills, and vice-versa. In the model presented in this study, I microfound the dynamics of skill acquisition upon these three concepts. The assumption of local cross-productivity [a_3] will first be employed in the full dynamic model, however. Below is a formalization of the first two concepts.

Definition 3.1. *We adopt the following convention.*

(a) A skill j is self-productive if skills produced at one stage augment the ones attained at later

stages, or formally,

$$\frac{\partial}{\partial a_j}\Delta h_j > 0$$

(b) A skill j satisfies dynamic complementarity if produced at one stage raise the productivity of human capital investment at subsequent stages, or formally,

$$\frac{\partial^2}{\partial a_j \partial h_{j,0}}\Delta h_j > 0$$

It follows promptly that the learning mechanism described in (2) is self-productive and satisfies dynamic complementarity in compliance with empirical findings [a_1] and [a_2] respectively.

It is also important to investigate the predictions of this setting for returns-to-education. The empirical literature on schooling returns has made an evolution from estimating constant to diminishing rates. Indeed, previously returns-to-schooling were thought of constant due to the seminal works of Mincer (1958, 1974). This seminal framework builds on a distinction between schooling and experience and typically assumes a psychological (hedonic) cost to education explicated in the utility function of agents. The resulting decision rules are simple and conveniently expressed as a regression prompting their almost five-decade long popularity in the empirical literature. Mincerian regression coefficients were commonly interpreted as *internal rates of return* (IRR) to education as explicated by Becker (2009). However, around the millennial shift evidence emerged calling into question the validity of Mincer models' theoretical construction and their corresponding interpretation. For instance, Katz and Autor (1999) and Heckman, Lochner and Todd (2006) reject the Mincerian functional forms. Employing a non-parametric approach, moreover, Heckman, Lochner and Todd (2008) find two deviations from key assumptions of the Mincer model – namely parallelism and linearity in log earnings which, in turn, are quantitatively important for estimating IRRs. Card (1999) and Heckman, Humphries and Veramendi (2018) find diminishing marginal returns to education. In agreement with the former piece of evidence, Heckman, Lochner and Todd (2006, 2008) and Heckman, Schmierer and Urzua (2010) find difference between yields of schooling on average and margin respectively. In the framework developed here, we can readily show that returns to learning a skill is increasing but at a diminishing rate.

Proposition 3.1. *Returns to learning a skill is positive but at a diminishing rate, that is,*

$$\frac{\partial y_j}{\partial a_j} > 0, \qquad \frac{\partial^2 y_j}{\partial a_j^2} < 0.$$

Proof. The proof is merely inserting h_j from (2) into the earnings function $y_j = w_j h_j \ell_j$ and taking the corresponding derivatives with respect to a_j. □

Hence, basing the model of this study on the dynamics of skill formation with focus on endowment effects, I am thusly able to replicate a sorting behavior and return structure consistent with these empirical observations. As such, in Proposition 3.1 we have shown that by assuming a learning scheme that is self-productive and dynamically complementary we obtain anincome structure that satisfies empirically verified properties, namely diminishing marginal returns and subsequent discrepancy between average and marginal returns to schooling as documented in the empirical literature listed with finding [1].

We assume now further that skill 1 is easier to learn, that is, $g_1(a) > g_2(a), a \in [0,1]$. For intance one can think of skill 1 as a more routine task, e.g. bookkeeping, while skill 2 is more non-routine e.g. software programming. The independent probability of skill obsolescence is given by $\Pr(w_j = 0) = \xi_j$. For ease of analysis we are going to use the following specification of the attention budget instead:

$$(a_1, a_2, p) \in [0, p] \times [0, 1-p] \times [0, 1] \tag{3}$$

with $p = p_1 \triangleq a_1 + \ell_1$, or equivalently, $1 - p = p_2 \triangleq a_2 + \ell_2$. In other words, p is the portion of the attention given to labor or learning of skill 1. Moreover, $\boldsymbol{p} \triangleq (p_1, p_2)$ is the agent's *job description*, which is described exhaustively through p. Hence, we will use the term job description interchangeably for p and $\boldsymbol{p}$. The job description can be thought of how much attention the agent gives to each task within their job profile: $p_1 = p$ is the attention share given to the routine task (e.g. bookkeeping) and $p_2 = 1 - p$ is the attention share devoted to the non-routine task (e.g. software-programming).

Thus the maximization problem becomes

$$\max_{a_1, a_2, p} \mathbb{E}(U(c)) = U(y_1 + y_2)(1-\xi_1)(1-\xi_2) + U(y_1)(1-\xi_1)\xi_2 + U(y_2)(1-\xi_2)\xi_1 + U(0)\xi_1\xi_2, \tag{4}$$

subject to (1), or equivalently (3). The utility function U increasing in consumption at a diminishing rate ($U' > 0$ and $U'' < 0$). We have indirectly assumed that the probability of the two tasks becoming obsolete is independent of one another.

From now on we adopt the following convention. We say that an agent *specializes in skill j* if the optimal choice of attention to learning for the other skill is zero, i.e. $a_i^* = 0$ for $i \neq j$. We say that an agent *focuses exhaustively on skill j* if the portion devoted to learning and labor with it adds up to the whole attention budget, i.e. $p_j^* = 1$.

In case the job description p is not fixed, in addition to (54) the following optimality condition holds for interior solutions:[1]

$$\frac{(1-\xi_1)w_1 h_1}{(1-\xi_2)w_2 h_2} = \frac{U'(y_1+y_2)(1-\xi_1) + U'(y_2)\xi_1}{U'(y_1+y_2)(1-\xi_2) + U'(y_1)\xi_2} \tag{5}$$

The equation above states that the optimal job description p^* is chosen such that the relative marginal gains from income in the skills equal their relative expected marginal utility. If the

[1]This is the result of the following first order condition: $\frac{\partial}{\partial p}\mathbb{E}(U(c)) = 0$.

left-hand side of (5) is larger than the right-hand side, then the agent's optimal decision is to focus exhaustively on skill 1 and vice versa.

We can derive the following interesting comparative statics for endowment profiles $h_{j,0}$ of the agents. The proposition states that optimal attention to learning and labor with skill j is increasing in the endowment in said skill $h_{j,0}$ provided that the wage rate of the other skill is low enough.

Proposition 3.2. *Wherever differentiable,* ceteris paribus *the following holds,*

$$\frac{\partial}{\partial h_{1,0}} a_1^* \geq 0, \qquad \frac{\partial}{\partial h_{1,0}} a_2^* \leq 0, \qquad \frac{\partial}{\partial h_{1,0}} p^* \geq 0 \tag{6}$$

if and only if

$$w_2 \leq \left(1 + \frac{\xi_2}{1-\xi_2} \cdot \frac{U''(y_1)}{U''(y_1+y_2)}\right) \frac{h_1}{h_2} w_1. \tag{7}$$

Remark 3.2.1. *Similar corresponding results can be derived for $h_{2,0}$ where inequalities in (6) are reversed and (7) is replaced by*

$$w_1 \leq \left(1 + \frac{\xi_1}{1-\xi_1} \cdot \frac{U''(y_2)}{U''(y_1+y_2)}\right) \frac{h_2}{h_1} w_2. \tag{8}$$

Observe that if $\ell_j^* = 0$ then $a_j^* = 0$. From Proposition (3.2) we can then find necessary and sufficient conditions for an agent to exhaustively focus on skill 2. Of course corresponding corollary can be derived for skill 1.

Corollary 3.2.1. *The agent focuses exhaustively on skill 2 if and only if*

$$h_{2,0} > \left(1 + \frac{\xi_2}{1-\xi_2}\right) \frac{w_1}{w_2} h_1 \mid_{\ell_2^* = a_2^* = 0} . \tag{9}$$

If skill 2 safe relative to skill 1 - i.e. $\xi_2 = 0$ - then the condition becomes,

$$w_2 h_{2,0} > w_1 h_{1,0}. \tag{10}$$

Proof. The follows directly from Proposition (3.2), where (9) provides the complement set characterized by (5) evaluated at $\ell_2^* = a_2^* = 0$. Finally, (10) follows from (9) when $\xi_2 = 0$. □

Having provided characteristics for exhaustive focus on skills, we move on to derive conditions for specialization upon skills, i.e. $a_j^* = 0$. We define the following *ceteris paribus lock-in thresholds*:

$$\underline{h}_{1,0} \equiv \sup_{a_1^*=0} h_{1,0}, \qquad \bar{h}_{1,0} \equiv \inf_{a_2^*=0} h_{1,0}$$

Since the optimal labor choice for skill j is given by

$$\ell_j^* = \frac{g_j(a_j^*)}{g_j'(a_j^*)}.$$

we can deduce that these thresholds have the following structure. Namely,

$$\underline{h}_{1,0} \text{ is given by } \frac{g_1(0)}{g_1'(0)} = p^*(\underline{h}_{1,0}) \quad \text{and} \quad \bar{h}_{1,0} \text{ is given by } \frac{g_2(0)}{g_2'(0)} = 1 - p^*(\bar{h}_{1,0}). \tag{11}$$

It is furthermore obvious that by construction $\underline{h}_{1,0} \leq \bar{h}_{1,0}$. For the remaining analysis we assume $\xi_2 \leq \xi_1$. In other words, the non-routine task 2 (e.g. software-programming) has a lower likelihood of becoming obsolete relative to the routine task 1 (i.e. bookkeeping), an assumption which is empirically motivated by the literature on RBTC [5].

Figures 1 and 2 go here.

In Figure 1 we can see these *ceteris paribus* endowment thresholds of lock-in for skill 1. To begin with, the figure illustrates the results of Proposition 3.2. We see that attention to learning and labor with skill 1 is increasing with respect to human capital endowment in said skill. Moreover, the spectrum divides agents into three types: specializers in skill 2, diversifiers and specializers in skill 1 corresponding to regions (I) , (II) and (III) respectively. An interesting dynamic, however, arises when skills are assumed to be safe, i.e. when the probabilities of obsolescence ξ_j approach zero. In that case, we can show that there are no diversifiers. This is depicted in Figure 2, where region (II) is nearly empty. We summarize the results in Proposition 3.3.

Proposition 3.3. *Let $\xi_2 \leq \xi_1$. Then $\bar{h}_{1,0} \to \underline{h}_{1,0}$ as $\xi_1 \to 0$. In other words, in absence of skill obsolescence agents purely specialize.*

Hence, direct consequence of Propositions (3.2) and (3.3) is the empirically verified results that agents sort into education based on realized returns (Heckman, Schmierer and Urzua, 2010; Carneiro, Heckman and Vytlacil, 2011). Consistent with findings of skill formation literature (esp. Cunha, Heckman and Schennach, 2010), Heckman, Humphries and Veramendi (2018) illustrate such selection bias and sorting gains in schooling which are akin to endowment effects in human capital for technology adoption (Comin and Hobijn, 2004; Foster and Rosenzweig, 2010). In other words, workers which have comparatively higher endowment in skills that are more difficult to acquire, tend to be more prone to work with new complex technology [3].

Recall that returns to human capital investment is increasing at a diminishing rate [1] which is in line with the overview of the empirical literature on returns to schooling (Card, 2001; Heckman, Humphries and Veramendi, 2018). Here we have shown that specialization is a consequence of skills being safe from obsolescence. Becker (1985) argues, however, that specialization is primarily prompted by increasing returns to education, while not considering skill obsolescence. Indeed, even in the model presented by Becker (1985) the main driver of specialization is the skills being safe. We summarize the result in the following corollary.

Corollary 3.3.1. *Increasing returns to human capital investment is not necessary for pure specialization.*

Another interesting dynamic arises if we consider intergenerational transmission of human capital. In the intergenerational interpretation of the setting here I adhere to viewing human capital stock also representing both sharpness of skills and extent of networks for the sector employing said skills. In other words each infinitely-lived dynasty transmits human capital in the form of both skills and the networks within the sector where the skills are utilized in line with finding [4]. Each generation in a dynasty thus teaches the next one their tools of the trade and embeds the coming generation in the web of relations that they have cultivated over the years. Imagine a situation where each generation faces the same problem as in (4), while the next generation inherits the new profile stock of human capital

$$(h_{1,n,i}, h_{2,n,i}) = (g_1(a^*_{1,n,i})h_{1,n-1,i}, g_1(a^*_{2,n-1,i})h_{2,n-1,i})$$

where $n \geq 1$ is the number of generation and $i \in \{1, \ldots, N\}$ is dynasty index and N is the population size. Figure 3 shows an example where there are diversifies in the population and the skill profiles are edged at the lock-in thresholds. The example can be seen as the new introduction of a new skill - h_2 - being recently monetized, why there is less endowment variation in the population than skill 1. After five iterations however, agents are starting to diverge and after ten generations the difference is steeper. Nevertheless, there always is the intermediate population of diversifiers acting as a bridge between the groups of specializers at the edges. On the other hand, in Figure 4 the agents do not perceive any risk of obsolescence for the skills, and so very quickly the population is cloven into two distinct types. These results suggest that - given the perceived threat of human capital obsolescence - differences in relative endowment among a population have profound evolutionary implications for the division of labor in society creating distinct dynastic human capital profiles [4].

Figures 3 and 4 go here.

Next, we extend the model to a dynamic scenario where the insights here carry over.

3.2 Full Model

In the full dynamic model we allow workers to save, which will drive the physical capital accumulation. Furthermore, they will invest in a continuum of skills ordered from most to least routine following the conceptualization in Autor, Levy and Murnane (2003) and Acemoglu and

Restrepo (2018*e*). First I will explicate the task and skill structure. Second, the production economy is described. Final goods producers employ tasks supplied by intermediate producers, who in turn wither use machines or labor for supplying tasks. Next, workers decision problem is stated, followed by elaborating on aggregation. Finally, an intergenerational interpretation of the setting is stated.

3.2.1 Tasks and Skills

The index framework is summarized in Figure 5. Time t is continuous. Tasks and skills j have a one-to-one correspondence and are ordered from most to least routine. The threshold I_t is marks available automation technology – i.e. the threshold below which it is possible to perform the tasks by machines. However, $\widetilde{I}_t$ is the threshold below which it is cheaper to perform the tasks by machines – provided that the technology is available. Then labor's exclusive threshold of factor production I_t^* is given by,

$$I_t^* = \min\{I_t, \widetilde{I}_t\}. \tag{12}$$

In other words, tasks below I_t^* are performed by machines, and those beyond it by labor. We will later on in the text specify conditions for which it holds that

$$I_t^* = I_t, \tag{13}$$

but for now we stipulate it as an assumption.

Figure 5 goes here.

Similarly we define the task index frontier N_t^* as:

$$N_t^* = \min\{N_t, \widetilde{N}_t\}, \tag{14}$$

where N_t is the index threshold above which technology is not available for labor to perform the corresponding tasks, while $\widetilde{N}_t$ is the threshold below which it is economically viable for the tasks to be performed – provided that the technology is available. Tasks produced in the economy lie in the interval $[N_t^* - 1, N_t^*]$. Tasks in the interval $[N_t^* - 1, I_t^*]$ are done by machines, while tasks in the interval $[I_t^*, N_t^*]$ are carried out by workers. In the setting of the static model developed earlier in Section 3.1, the more routine skill 1 (bookkeeping) would then appear after I_t^*, but before then non-routine skill 2 (software development), which in turn precedes N_t^*. We will later on in the text specify conditions for which it holds that

$$N_t^* = \widetilde{N}_t, \tag{15}$$

but for now we stipulate it as an assumption. This framework is very similar to Acemoglu and Restrepo (2018e), with the important difference that there exists an interval with unadopted new tasks. Theses tasks are deemed not to be economically viable. The reason is lack of human capital in those particular tasks, as will be illustrated later. The arrival of new technology I_t and available index frontier N_t – new tasks which replaces old tasks – are exogenous and given by the following jump processes:

$$dI_t = dJ_I(t), \text{ where } dJ_I(t) \sim Poi(\lambda) \cdot f_{\Delta_I}(s), f_{\Delta_I} : [0, n_t^*] \to \mathbb{R}_+ \tag{16}$$

$$dN_t = dJ_N(t), \text{ where } dJ_N(t) \sim Poi(\lambda) \cdot f_{\Delta_N}(s), f_{\Delta_N} : [0, 1 - n_t^*] \to \mathbb{R}_+ \tag{17}$$

where $n_t^* = N_t^* - I_t^*$. Both processes are Poisson arrivals with the same intensity $\lambda > 0$ but different jump-size densities f_{Δ_I} and f_{Δ_N}.

The rates of arrival, can be endogenized by assuming a pool of scientists being allocated between research and development on automation and new-task creation as illustrated in Acemoglu and Restrepo (2018e). However, as explained farther in this study, given the slower rate of accumulation in human vis-à-vis physical capital, it is the institutional rate of adoption that is decisive. Moreover, Acemoglu and Restrepo (2018e) show that for balanced growth paths, equal arrival rate of both technologies is a necessary condition. Hence, by assuming an exogenous rate of arrival, we abstract away an immaterial mechanism while rigging the growth path in favor of being balanced. Thus, if unbalanced growth is to follow yet, in less favorable scenarios it will do so still.

For each task $j \in [I_t^*, N_t^*]$ there exists a corresponding skill j for which workers $i \in [0, 1]$ at time t have human capital stock h_{ijt}. This stock augments their corresponding labor supply ℓ_{ijt} when earning income $y_{ijt}^L = w_{jt} h_{ijt} \ell_{ijt}$ where w_{jt} is the wage rate for task j at time t.

3.2.2 Producers

Final-Goods Producers: A competitive, risk-neutral and representative final goods producer homothetically aggregates a continuum of intermediate tasks indexed by $j \in [N_t^* - 1, N_t^*]$,

$$Y_t = \left(\int_{N_t^*-1}^{N_t^*} y_{jt}^{\frac{\varepsilon-1}{\varepsilon}} \, dj \right)^{\frac{\varepsilon}{\varepsilon-1}} \tag{18}$$

where $\varepsilon > 0$ is the technical elasticity of substitution between tasks. Since we are discussing division of labor in society, I follow Becker and Murphy (1992), Acemoglu (1998) and Autor, Levy and Murnane (2003) in assuming tasks are complementary to varying degrees. This is in line with the empirical findings of Dinopoulos et al. (2011) who find that at the aggregate high- and low-skilled workers are gross complements. Moreover, the technology-skill complementarity is well-documented in the empirical literature.[2] Formally, this assumptions is expressed as follows:

[2]See for instance assumption A2 in Autor, Levy and Murnane (2003) for how routine and non-routine tasks are assumed to be imperfect substitutes. For more examples of technology-skill complementarity see Goldin and Katz (1998, 2009); Krusell et al. (2000); Brynjolfsson and Hitt (2003); Caselli and Coleman (2006).

Assumption 1.

Tasks are imperfect substitutes but gross complements of one another, i.e. $0 < \varepsilon < 1$.

Cost minimization yields then the following demand for intermediate task j:

$$y_{jt}^D = \left(\frac{p_{jt}}{P_t}\right)^{-\varepsilon} Y_t, \text{ where } P_t \equiv \left(\int_{N_t^*-1}^{N_t^*} p_{jt}^{1-\varepsilon} dj\right)^{\frac{1}{1-\varepsilon}}, \tag{19}$$

where y_{jt}^D is the demand for task j at time and p_{jt} is its price.

Task Producers: We assume that there is a one-to-one correspondence between tasks and skills. Following Acemoglu and Restrepo (2018*a*,*e*), I assume that agents use technology $\vartheta_L(j)$ together with their human-capital-augmented labor supply for each task $h_{jt}\ell_{jt}$

$$h_{jt}\ell_{jt} \equiv \int_0^1 h_{ijt}\ell_{ijt} di, \tag{20}$$

to produce the corresponding task $j \in [N_t^* - 1, N_t^*]$ as $y_{jt}^S = \vartheta_L(j)h_{jt}\ell_{jt}$. Nevertheless, for tasks in the interval $[N_t^* - 1, I]$, machines can also produce the task as perfect substitutes using the technology $\vartheta_M(j)$ and capital devoted to that task k_{jt} in accordance with $y_{jt}^S = \vartheta_M(j)k_{jt}$.[3] Hence, we have the following production function:

$$y_{jt}^S = \begin{cases} \vartheta_M(j)k_{jt} + \vartheta_L(j)h_{jt}\ell_{jt} & \text{for } j \in [N_t^* - 1, I_t] \\ \vartheta_L(j)h_{jt}\ell_{jt} & \text{for } j \in [I_t, N_t^*] \end{cases} \tag{21}$$

Just as Acemoglu and Restrepo (2018*a*,*e*) assume that $\vartheta_L(j)$, $\vartheta_M(j)$ and $\vartheta_L(j)/\vartheta_M(j)$ are increasing in j, where the last assumption implies that labor has comparative advantage in the production of high-indexed tasks compared to machines. Observe that for tasks $j \in [N_t^* - 1, I]$, labor and machines are perfect substitutes so production is done with the cheaper factor. By setting demand for tasks y_{jt}^D in (19) equal to their supply y_{jt}^S provided by (21) we can derive expressions for their prices.

Lemma 3.1. *The price of tasks are given by*

$$p_{jt} = \begin{cases} P_t\left(\frac{Y_t}{\vartheta_M(j)k_{jt}}\right)^{\frac{1}{\varepsilon}} & \textit{for } j \in [N_t^* - 1, I_t^*] \\ P_t\left(\frac{Y_t}{\vartheta_L(j)h_{jt}\ell_{jt}}\right)^{\frac{1}{\varepsilon}} & \textit{for } j \in [I_t^*, N_t^*] \end{cases}$$

Finally, we explicate a subtle assumption under which we have operated so far,

Assumption 2.

Rental rate of capital is the same for all the tasks $j \in [N_t^* - 1, I_t^*] : r_{jt} = r_t$.

[3] See Autor and Dorn (2013) for evidence on capital and labor being gross substitutes. See León-Ledesma, McAdam and Willman (2010) and references therein for evidence to the contrary.

This assumption has allowed us to abstract from the investment decisions. This can be seen as all agents investing in the same stock market where a no-arbitrage condition has rendered all indexes fiscally equivalent. Thereby, Assumption 2 will yield the relative allocation of the capital stock. Observe that the following capital market clearing condition holds:

$$K_t = X_t \text{ where } K_t \equiv \int_{N_t^*-1}^{I_t^*} k_{jt} dj \text{ and } X_t \equiv \int_0^1 x_{it} di. \tag{22}$$

where x_{it} is the financial assets of agent i at time t, X_t its corresponding aggregate stock, and K_t the aggregate physical capital stock. The economy's market clearing condition,

$$r_t X_t + Y_t^L = Y_t \text{ where } Y_t^L \equiv \int_0^1 \int_{I_t^*}^{N_t^*} y_{ijt}^L dj di, \tag{23}$$

or equivalently

$$S_t + p_t C_t = Y_t \text{ where } C_t \equiv \int_0^1 c_{it} di. \tag{24}$$

and S_t are the aggregate savings at time t, c_{it} individual worker i's consumption at time t at price p_t^c plus its aggregate stock C_t. Assuming perfect competition among risk-neutral task producers, we can then prove the following proposition.

Proposition 3.4. *Let the market be perfectly competitive with task producers having production function (21) and operating under Assumption 2. Then rental rate of capital and the wage rates are given by*

$$r_t = P_t \left(\frac{Y_t \cdot \int_{N_t^*-1}^{I_t^*} \vartheta_M^{\varepsilon-1}(j) dj}{X_t} \right)^{\frac{1}{\varepsilon}} \text{ and } w_{jt} = P_t \left(\frac{Y_t}{h_{jt} \ell_{jt}} \right)^{\frac{1}{\epsilon}} \vartheta_L^{\frac{\varepsilon-1}{\varepsilon}}(j) \tag{25}$$

for $j \in [N_t^* - 1, I_t^*]$ and $j \in [I_t^*, N_t^*]$ respectively. Moreover, capital dedicated to task $j \in [N_t^* - 1, I_t^*]$ is given by

$$k_{jt} = \mathcal{P}_{jt} X_t \text{ where } \mathcal{P}_{jt} \equiv \frac{\vartheta_M^{\varepsilon-1}(j)}{\int_{N_t^*-1}^{I_t^*} \vartheta_M^{\varepsilon-1}(j) dj}, \tag{26}$$

where $\mathcal{P}_{jt}$ is a density function describing the distribution of capital among tasks $j \in [N_t^* - 1, I^*]$.

We derive the following corollaries of Proposition (3.4):

Corollary 3.4.1. *The price of tasks produced by capital $j \in [N_t^* - 1, I_t^*]$ is given by*

$$p_{jt} = \frac{P_t}{\vartheta_M(j)} \left(\frac{Y_t}{X_t} \int_{N_t^*-1}^{I_t^*} \vartheta_M^{\varepsilon-1}(j) dj \right)^{\frac{1}{\varepsilon}}$$

Proof. The proof follows promptly from inserting (26) into the expression given by Lemma 3.1. ∎

Corollary 3.4.2. *The portion of aggregate capital invested in the skill-index interval* $[j', j]$, $j' \leq j$ *and* $j', j \in [N_t^*, I_t^*]$ *at time* t *is dubbed* $\mathcal{P}_{[j',j],t}$ *and is given by*

$$\mathcal{P}_{[j',j],t} = \frac{\int_{j'}^{j} \vartheta_M^{\varepsilon-1}(z)dz}{\int_{N_t^*-1}^{I_t^*} \vartheta_M^{\varepsilon-1}(z)dz}.$$

Proof. This corollary is a direct consequence of (26). □

Tasks will be produced by machines if and only they are cheaper, that is,

$$p_{jt}^M < p_{jt}^L$$

which by Lemma 3.1 yields

$$\frac{\vartheta_L(j)}{\vartheta_M(j)} < \frac{k_{jt}}{h_{jt}\ell_{jt}}.$$

Hence, by (26) we have that

$$h_{jt}\ell_{jt} < \frac{\vartheta_M^{\varepsilon}(j)}{\vartheta_L(j)} \cdot \frac{X_t}{\int_{N_t^*-1}^{I_t^*} \vartheta_M^{\varepsilon-1}(j)dj}. \tag{27}$$

Since by assumption $\vartheta_L(j)/\vartheta_M(j)$ is increasing in j and tasks are complementary $(0 < \varepsilon < 1)$, the right-hand side above is decreasing in the task index. Thereby, there exists a threshold $\tilde{I}_t$ given by

$$h_{\tilde{I}_t,t}\ell_{\tilde{I},t} = \frac{\vartheta_M^{\varepsilon}(\tilde{I}_t)}{\vartheta_L(\tilde{I}_t)} \cdot \frac{X_t}{\int_{N_t^*-1}^{\tilde{I}_t} \vartheta_M^{\varepsilon-1}(j)dj}. \tag{28}$$

such that for all tasks $j \in [N_t^* - 1, \tilde{I}_t]$ are produced by machines, provided that the technology exists, i.e. $\tilde{I}_t \geq I_t$. Recall that $I_t^* = \min\{I_t, \tilde{I}_t\}$ where I_t is the automation threshold. By (28), we have the following condition for new automation technology to be adopted:

$$h_{I_t,t}\ell_{I_t,t} < \frac{\vartheta_M^{\varepsilon}(I_t)}{\vartheta_L(I_t)} \cdot \frac{X_t}{\int_{N_t^*-1}^{I_t} \vartheta_M^{\varepsilon-1}(j)dj}. \tag{29}$$

We can see here that a decrease in labor productivity ϑ_L will increase the comparative advantage of physical capital and hence ease automatization – i.e. routine-biased technological change [5]. Moreover, when automation occurs, productivity increases, but as some significant stock of human capital of agents adept at routine tasks become obsolete, polarization follows [6].

To explore whether new tasks are adopted, we need to adopt some framework for intellectual property. As new tasks become available N_t^*, old tasks $N_t^* - 1$ are threatened with obsolescence. The owners to the old technology's copy-rights need to be compensated if the new tasks pose any infringement. Ponder a situation where the owners need to be fully compensated. Then the new technology will not shrink the economy, that is,

$$y_{N_t^*,t} > y_{N_t^*-1,t} \tag{30}$$

which by (21) and (26) yields

$$h_{N_t^*,t}\ell_{N_t^*,t} > \frac{\vartheta_M^{\varepsilon}(N_t^*-1)}{\vartheta_L(N_t^*)} \cdot \frac{X_t}{\int_{N_t^*-1}^{I_t^*} \vartheta_M^{\varepsilon-1}(j)dj}. \tag{31}$$

Now consider a situation where there are no laws protecting intellectual properties. In such a case new technology is adopted if it is cheaper, risking shrinkage of the economy. In other words,

$$p_{N_t^*,t} < p_{N_t^*-1,t} \tag{32}$$

which by Lemma 3.1 and (26) yield once again the condition in (31). This is of course a result of assuming perfect competition. As the two extreme frameworks of intellectual property yield the same adoption conditions, all the intermediate cases will as well.

Comparing conditions for adoption of automation (29) and new tasks (31) we see that capital accumulation - i.e. an increase in X_t - has inverse effects. Indeed, capital accumulation increases the possibility of automation adoption by reducing the price of capital, while it has the inverse effect on the adoption of new tasks through the same mechanism. Moreover, new technology – automation or new tasks – is adopted if it offers higher productivity. Furthermore, as we will see, the faster rate of physical capital accumulation will entail that automation technology is more readily adopted than new tasks leading to routine-biased technological change [5], and falling labor share of production [7] .

Finally, we can employ (25) in order to derive the relative skill premiums for indexes $k > j$,

$$\frac{w_{kt}}{w_{jt}} = \left[\underbrace{\left(\frac{\vartheta_L(j)}{\vartheta_L(k)}\right)^{1-\varepsilon}}_{\text{Productivity discount}} \cdot \underbrace{\frac{h_{jt}\ell_{jt}}{h_{kt}\ell_{kt}}}_{\text{Scarcity premium}} \right]^{\frac{1}{\varepsilon}}. \tag{33}$$

Since labor has comparative advantage in higher-indexed skills ($\vartheta_L(j)$ is increasing in j) and tasks are imperfect complements ($0 < \varepsilon < 1$), the productivity premium is less than one. Hence, a more appropriate word is productivity *discount*, rather than premium. This is related to the so called "cost disease" outlined by Baumol (1967), a phenomenon which is incidentally a major theme in the analysis of Aghion, Jones and Jones (2019) on the impact of artificial intelligence on economic growth. Aghion, Jones and Jones (2019, p. 241) define Baumol's cost disease as the phenomenon where production and "economic growth may be constrained not by what we do well but rather by what is essential and yet hard to improve." In the setting here, the cost disease translates into the low-productivity tasks becoming increasingly important in the economy since they are still required as gross complements. These tasks are essential yet hard to improve, so as long as their producers are not too abundant they will earn a higher wage. More generally, whether wages of higher-indexed tasks are higher than lower-indexed ones depends on the relative scarcity of their human-capital-augmented labor supply. In other words, the wage rate of higher-indexed tasks is larger if and only if their corresponding human-capital-augmented labor supply is much scarcer - or more precisely - scarcer by more than a factor of the productivity premium.

3.2.3 Workers

We move on to a dynamic model of skill diversification with a continuum of skills where agents $i \in [0,1]$ face the following problem

$$\max_{a_{ijt},\ell_{ijt},c_{it}} \mathbb{E}_0 \int_0^\infty U(c_{it}) e^{-\rho t} dt \tag{34}$$

where $j \in [N_t^* - 1, N_t^*]$ is the index of skills, N_t^* is the technology index frontier and the rest as discussed earlier. Income from skill j is given by,

$$y_{ijt}^L = w_{jt} h_{ijt} \ell_{ijt}, \tag{35}$$

where w_{jt} is the wage rate for task j at time t and h_{ijt} is the agent i's stock of human capital in skill j at time t. An agent's assets x_{it} is developed in accordance with the following law of motion:

$$dx_{it} = (r_t x_{it} + \int_{I_t^*}^{N_t^*} y_{ijt}^L dj - p_t^c c_{it})dt + \sigma_i(x_{it}, \boldsymbol{h}_{it}) dB_{it} \tag{36}$$

where r_t is the rental rate of capital, I_t^* is the task index threshold above which labor is the exclusive factor of production, p_t^c is the price of consumption, $\sigma_i(x_{it}, \boldsymbol{h}_{it})$ is an idiosyncratic volatility function, $\boldsymbol{h}_{it} = (h_{ijt})_{j \in [I_t^*, N_t^*]}$ is the *stock-density* functional of human capital,[4] and dB_{it} is an idiosyncratic Brownian motion. We dub $\boldsymbol{h}_{it}$ stock-density rather than density as we assume that the individual's aggregate stock of human capital h_{it} does not have to be 1. In other words,

$$h_{it} = \int_{I_t^*}^{N_t^*} h_{ijt} dj \geq 0. \tag{37}$$

We can then express the stock-density functional as the product of individual's aggregated stock h_i and some proper density function f_{h_i},

$$\boldsymbol{h}_{it} = h_{it} \left(f_{h_{it}}(j) \right)_{j \in [I_t^*, N_t^*]}, \tag{38}$$

where $f_{h_{it}}(j) = h_{ijt}/h_{it}$, so that $\int_{I_t^*}^{N_t^*} f_{h_{it}}(j) dj = 1$. Initial stock profile $\boldsymbol{h}_{i0}$ is observed by the agent. We assume the following distributional structure for the density, $f_{h_{i0}}(j) \sim e^{-\gamma_{h_i} j}$ depicted in Figure 6. As such all agents have more stock in low-indexed skills than high-indexed,

[4] Indeed the agent might have stock of human capital for $j \in [0, I_t^*]$ but that is irrelevant to the decision of the agent and hence to the production.

but those with lower γ_{h_i} have comparatively more stock in the latter. However, we assume that the endowment in aggregate stock of human capital is the same for agents in the population, i.e. $h_{i0} = h_0 > 0$. The heterogeneity structure for endowment in human capital has the form $1/\gamma_{h_i} \sim \text{Exp}(\varrho)$, i.e. an exponential distribution with expectation $1/\varrho$. This guarantees that there are more agents with predominant initial stock in routine skills, than those with comparative abundance in non-routine ones.

Figure 6 goes here.

The law of motion for agent's human capital is given by:

$$\dot{h}_{ijt} = g_j(a_{ijt})h_{ijt}, \tag{39}$$

where the learning function g_j satisfies the following conditions,

$$\text{(a) } g_j(0) = 0, \qquad \text{(b) } \frac{\partial}{\partial j} g_j < 0, \qquad \text{(c) } \frac{\partial}{\partial a_j} g_j > 0, \qquad \text{(d) } \frac{\partial^2}{\partial a_j^2} g_j < 0. \tag{40}$$

The requirement (40a) indicates that no learning leaves the stock of human capital unchanged and (40b) states that higher-indexed skills are more difficult to learn. The conditions (40c) and (40d) are recurrent and indicate increasing learning for any given level of attention a_{ijt} at a diminishing rate. We modify Definition (3.1) in the following manner for continuous-time framework.

Definition 3.2. *We adopt the following convention.*

(a) A skill j is self-productive if skills produced at one stage augment the ones attained at later stages, or formally,

$$\frac{\partial}{\partial a_j} \dot{h}_j > 0$$

(b) A skill j satisfies dynamic complementarity if produced at one stage raise the productivity of human capital investment at subsequent stages, or formally,

$$\frac{\partial^2}{\partial a_j \partial h_j} \dot{h}_j > 0$$

Once again, it follows promptly that the learning mechanism described in (40) is self-productive [a_1] and satisfies dynamic complementarity [a_2]. We can then readily show that returns to learning a skill is increasing but at a diminishing rate.

Proposition 3.5. *Returns to learning a skill is increasing but at a diminishing rate, that is,*

$$\frac{\partial y_j^L}{\partial a_j} > 0, \qquad\qquad \frac{\partial^2 y_j^L}{\partial a_j^2} < 0.$$

Proof. The proof is replacing h_{ijt} with $h_{ijt^-}e^{\dot{h}_{jt}}$ into the earnings function (35), then employing (40) and finally taking the corresponding derivatives with respect to a_j. □

Hence we have again shown that by assuming a learning scheme that is self-productive and dynamically complementary we obtain a learning function that satisfies empirically verified properties, namely diminishing marginal returns and subsequent discrepancy between average and marginal returns to schooling [1]. Also due to the *ex-ante* and *ex-post* heterogeneities in relative skill stock and wealth shocks respectively, we will have corresponding heterogeneity in returns to education across skill profile and time [2].

Finally, the attention budget is given by,

$$\int_{I_t^*}^{N_t^*} (a_{ijt} + \ell_{ijt})dj = 1. \tag{41}$$

The optimization problem (34) subject to (36), (39), (16), (17) and (41) is not in general tractably solvable. Indeed we need to choose the Bernoulli utility U, the learning functionals g_j in a way so as guarantee the ability to solve the ensuing Hamilton-Jacobi-Bellman (HJB) equation and the arising functional partial differential equations. We therefore assume the following structure:

Assumption 3.

$$U(c) = c - \frac{1}{2}\alpha c^2 \textit{ where } 0 < \alpha \ll 1, \quad \textit{and} \quad g_j(a) = e^{-\gamma_g j}(a - \frac{1}{2}\xi a^2) \textit{ where } \gamma_g > 0 \textit{ and } 0 \ll \xi < 1.$$

Assuming a linear-quadratic Bernoulli utility (felicity) function removes some quantitative nuance from the agent's saving behavior but allows us to get tractability elsewhere, namely the decisions on human capital investment. Nevertheless, this functional form imposes some limitations. It indeed implies increasing absolute risk aversion and has a point of satiation. Nevertheless, this utility could be seen as a second order Taylor approximation of the following logarithmic utility function,

$$\widetilde{U}(c) = \frac{2}{\alpha}\ln(1 + \frac{1}{2}\alpha c) = c - \frac{1}{2}\alpha c^2 + o(c^3).$$

Since the Taylor expansion of the logarithmic function has unity as radius of convergence, then $U(c)$ in Assumption 3 is a good approximation of $\widetilde{U}(c)$ as long as

$$0 < c < 2/\alpha.$$

Hence, by choosing α small enough we guarantee that this approximation is well-motivated and at the same time hinder the agent being able to reach the point of consumption satiation. Indeed it is due to this line of argument that the constrain on α is expressed as $0 < \alpha \ll 1$, indicating a very small positive number. The linear quadratic utility function allows us to employ a linear-quadratic *ansatz* for the optimal current-value function for the corresponding HJB equation. Moreover, since it is an approximation of logarithmic utility, it renders income and substitution effects to cancel out. However, it limits us to the case of near-unit constant relative risk-aversion. Nevertheless, for a more general utility structure numerical methods such as those developed by Ahn et al. (2018), Kaplan, Moll and Violante (2018) and Nuño and Moll (2018).

The functional form of g_j satisfies self-productivity. It is adjusted so that higher-indexed skills are more difficult to learn. More attention to learning a_j, however, leads to more stock of human capital h_j for a particular skill j at a diminishing rate. As mentioned earlier in the static model, a consequence of such functional form is the empirically plausible observation of increasing returns to schooling at a diminishing rate (Card, 2001; Heckman, Humphries and Veramendi, 2018). We interpret γ_g as the state of institutional education in the economy. Higher γ_g, indicates easier overall learning due to better facilities and pedagogical capabilities of teachers, instructors and communicators of knowledge in general, and vice versa. Observe however, that as $\partial^2 g_j / \partial \gamma_g^2 > 0$, more advanced institutional quality in the education system has larger effect on higher index skills – that is, those which are more difficult to obtain.

We divide workers into two types: diversifiers and specializers. A diversifier devotes some attention to learning every skill j corresponding to tasks that currently is being done by labor, i.e. $j \in [I_t^*, N_t^*]$. A specializer on the other hand, only learns skills corresponding to a subset $\mathcal{D} \subseteq [I_t^*, N_t^*]$ of labor-performed tasks. The magnitude of the spectrum of tasks performed by a specializer is given by $\nu(\mathcal{D})$, where $\nu : \mathcal{F}([I_t^*, N_t^*]) \mapsto \mathbb{R}_+$ is a corresponding measure function such that $\nu(\emptyset) = 0$ and $\nu([I_t^*, N_t^*]) = n_t^*$ with $\mathcal{F}([I_t^*, N_t^*])$ being the smallest possible sigma-algebra defined on $[I_t^*, N_t^*]$. The *average human-capital-augmented wage rate* for agent i at time t is defined as:

$$\overline{wh}_{it}^{\mathcal{D}} \equiv \frac{1}{\nu(\mathcal{D})} \int_{\mathcal{D}} w_{j,t} h_{ijt} dj. \tag{42}$$

The set $\mathcal{D}$ is dubbed the *set of manageable skills* and is characterized by

$$\begin{aligned} &w_{jt} h_{ijt} F_{I_t + \Delta_{I_t}}(j) = w_{j't} h_{ij't} F_{I_t + \Delta_{I_t}}(j') \text{ for all } j, j' \in \mathcal{D}, \\ &\qquad \text{while } w_{jt} h_{ijt} F_{I_t + \Delta_{I_t}}(j) > w_{kt} h_{ikt} F_{I_t + \Delta_{I_t}}(k), \text{ for all } j \in \mathcal{D} \text{ and for all } k \in \mathcal{D}^c. \end{aligned} \tag{43}$$

where $\mathcal{D}^c = [I_t^*, N_t^*] \setminus \mathcal{D}$. This condition is analogous to (7) and (8) in the static model. Observe that the quotas between second-order derivatives of utility cancel out due to the linear-quadratic utility structure. The condition (43) states that the expected income gain of the manageable skills in $\mathcal{D}$ are the same and more than the other skills in $\mathcal{D}^c$. We now make explicit the risk and volatility structure under non-zero probability of skill obsolescence.

Assumption 4.

Jump intensity $\lambda > 0$, *and asset risk given by* $\sigma_i(x_{it}, \boldsymbol{h}_{it}) = \sigma\sqrt{2x\left(\widetilde{Q} + \frac{D_{1t}}{2D_{2t}}\right)}$,

where,

$$\widetilde{Q} = x_{it} + \int_{I_t^*}^{N_t^*} D_{ijt}\sqrt{2h_{ijt}}dj, \tag{44}$$

and D_{ijt}*,* D_{1t} *and* D_{2t} *are given by,*

$$D_{ijt} = \sqrt{\frac{2e^{\gamma_g j}\overline{wh}_{it}w_{ijt}F_{\Delta_{I_t}}(j)}{4\xi(r_t + \sigma^2 - \lambda\mathcal{P}_{\Delta_{N_t^*}}) - e^{-\gamma_g j}}}, \tag{45}$$

$$D_{2t} = \frac{\alpha(\rho - 2(r_t + \sigma^2 - \lambda\mathcal{P}_{\Delta_{N_t^*}}))}{6\,(p_t^c)^2} \quad \textit{and,} \tag{46}$$

$$D_{1t} = \frac{\left(\rho - 2(r_t + \sigma^2 - \lambda\mathcal{P}_{\Delta_{N_t^*}})\right)\left[1 + \frac{\alpha\overline{wh}_{it}\left(1 - \frac{1}{\xi}n_t^*\right)}{p_t^c}\right]}{p_t^c\left(4\rho - 5(r_t + \sigma^2 - \lambda\mathcal{P}_{\Delta_{N_t^*}})\right)}, \tag{47}$$

with

$$\mathcal{P}_{\Delta_{N_t^*}} \equiv \mathcal{P}_{[N_t^*-1, N_t^*-1+\mathbb{E}\,\Delta_{N_t^*}],t}. \tag{48}$$

The volatility structure above is a qualified geometric diffusion. Just as the plain geometric diffusion process, however, the volatility in this setting approaches zero as financial capital stock x diminishes indefinitely, which in turn guaranties non-negative values on assets. The particular qualifications here are in place mainly so as to achieve tractability, but nevertheless do not entail any unusual consequences. For instance, one implication of the qualifications are that agents with larger stocks of human will have higher volatility in their portfolios, which is correlationally sound. Nevertheless, the volatility structure expressed in Assumption 4 is admittedly convoluted. Removing asset volatility ($\sigma = 0$), would, however, not get rid of these expressions in the optimal decision rules which are derived in Proposition 3.6 below, though they would appear only as a result which would be theoretically preferable.

Proposition 3.6. *Under assumptions 3 and 4, the agent with manageable human-capital-augmented wage rate facing the problem in (34) subject to (16), (17), (36), (39), and (41), acts in accordance with the following decision rules:*

$$a_{ijt}^* = \frac{1}{\xi}\left(1 - \sqrt{(4e^{\gamma_g j}\xi(r_t + \sigma^2 - \lambda\mathcal{P}_{\Delta_{N_t^*}}) - 1)\frac{w_{jt}h_{ijt}F_{\Delta_{I_t}}(j)}{\overline{wh}_{it}}}\right), \qquad \ell_{ijt}^* = \frac{1}{n_t^*} - a_{ijt}^* \tag{49}$$

$$c_{it}^* = \frac{1}{\alpha}\left(1 - D_{1it} - D_{2t}\widetilde{Q}\right), \tag{50}$$

where $\widetilde{Q}$ is the same as in (44), and D_{ijt}, D_{1t} and D_{2t} are given by (45), (46) and (47).

A consequence of Proposition 3.6 is that learning decisions are independent of asset stock x_t. Moreover, consumption is increasing in financial assets x_t if

$$r_t > \frac{1}{2}\rho + \lambda \mathcal{P}_{\Delta_{N_t^*}} - \sigma^2,$$

by (46), i.e. if rate of interest is large enough. Indeed it needs of set both rate of time preference and expected new task creation net asset volatility. In Sections 3.3 and 3.4 I will illustrate how the decisions of workers will lead to labor immiseration and job market polarization. However, first for conceptual completeness, I also provide an overlapping generations interpretation for the setting.

3.2.4 Intergenerational Transmission of Human Capital

We now extend our analysis to overlapping generations (OLG) of workers, who face a hazard of dying and are replaced by new generations, as in Yaari (1965) and Blanchard (1985). In the intergenerational interpretation of the setting here I adhere to viewing human capital stock also representing both sharpness of skills and extent of networks for the sector employing said skills. In other word, each infinitely-lived dynasty transmits human capital from one generation to another in the form of both skills and the networks within the sector where the skills are utilized. With the same intensity of a Poisson arrival of η, old workers die and new workers are born so that the total population is stable, and normalized to 1. The results can be extended to the case of non-constant death hazard over the life cycle as in Calvo and Obstfeld (1988). We further assume that the wealth of deceased workers is transmitted to the surviving generations though bequests or perfect annuity markets. Then, the wealth distribution will be unchanged. The line of argument so far is the same as Itskhoki and Moll (2019). For human capital we assume that the density f_h is passed on exactly as it is while the only a fraction υ of the aggregate stock reaches the next generation. This problem is equivalent to starting with a fraction υ of the initial human capital stock h_0, and thus a normalization to one will make the problem once again equivalent to one with infinitely lived agents with qualified discount factors $\rho + \eta$. A consequence of this interpretation is the perpetuation of agent-types (low- and high-indexed) through dynastic human capital [4].

3.3 Balanced Growth Paths and Labor Immiseration

In this section I will characterize the balanced growth path. In order to show the inevitability of labor immiseration, we need to show two things; that all tasks will ultimately be automated and that adoption of new tasks occur at a slower rate than rate of automation. Indeed when aggregate output Y_t grows at a constant rate, since learning higher-indexed skills becomes increasingly difficult ($\frac{\partial}{\partial j} g_j < 0$), all tasks will eventually be automated at a rate faster than adoption of new tasks. However, first we need to introduce (local) cross-productivity in skill formation [a_3] to have ripe conditions for balanced growth paths to emerge.

Definition 3.3. *Let $\mathcal{D}_{it}$ be the convex set of dominating and manageable skills for agent i at time t. We say that the worker is cross-productive with distance $\varsigma > 0$ in their skill if there exists a ball with radius ς at the boundary points of $\mathcal{D}_{it}$ where the stock-density function $\boldsymbol{h}_{it}$ extends along its analytic continuation.*

Definition 3.3 entails that workers have some latent stock of human capital close to their boundary skills, why we dub it *local* cross-productivity. When workers are cross-productive with distance $\varsigma > 0$, they do not end up with zero labor income if I_t^* extends beyond the skills that some worker is actively accumulating at the time. Moreover, the fact that stock-density function $\boldsymbol{h}_{it}$ extends along its analytic continuation ensures ordinal constancy of skill profiles. Crossproductivity articulated in Definition 3.3 entails that the stock-density of human capital $\boldsymbol{h}_{it}$ extend along its analytic continuation. This continuation for the low-index worker is decreasing in j. In other words, they have more stock in routine stocks. Similarly, the continuation of the stock-density is increasing for the high-index why they have more stock in non-routine skills. In other words, if the worker has more stock in routine tasks, it will continue to do so, and vice versa. For the rest of the discussion we assume that workers are cross-productive with distance $\varsigma > 0$ which is large enough so that $N_t^* + \varsigma > N_t + dN_t$. This is a technical assumption and guarantees that agents working at the edge of innovation technology N_t^* have latent stock of human capital beyond that threshold, so that if and when such technology is available some workers will have some stock to produce these tasks. This assumption therefore guarantees that there always exists some stock for skills to be used in production, giving workers an actual chance at competing with machines, and by extension a conceivable balanced growth path without labor immiseration. It is important to note that workers focus their attention to learning new tasks only once they are invented, and only if optimal. Moreover, workers do not accept lower wages since they can rely on capital income from their savings. In fact, if all tasks are automated, the population will solely rely on capital income, which of course is heterogeneous and unequal.

Any conception of a balanced growth path would require output Y_t to grow, at a constant rate say $\chi^* > 0$. Without loss of generality, we set consumption as numéraire, i.e. $p_t^c \equiv 1$. Then by the market clearing condition (24), aggregate consumption C_t also grows at the same rate $\chi^* > 0$ as aggregate output Y_t.[5] Moreover, by (50) and (44), aggregate stock of capital X_t also grows at this same rate. By the alternative formulation of the market clearing condition (23) it follows that the interest rate is constant $r_t \equiv r^* > 0$, which entails that the price of a task j performed by machines are also constant $p_{jt}^M = p_j^* = \frac{r^*}{\vartheta_M(j)}$. Moreover, by (23) either aggregate labor income Y_t^L grows at the same rate as output Y_t or it holds that $Y_t^L = 0$ for some t and onward. Below follows a discussion outlining why on any balanced growth path the latter holds in line with falling labor shares of income [7]. The proof is a *reductio ad absurdum*.

For any skill j aggregate labor income is given $y_{jt}^L = w_{jt}h_{jt}\ell_{jt}$. For the growth rate of y_{jt}^L

[5] On the balanced growth path per-capita consumption grows at a constant rate giving more and more weight to the negative quadratic term and turning marginal utility negative at some point for agents where $c_t > 1/\alpha$.Despite $0 < \alpha \ll 1$ being set very small, this is inevitable. There is a simple way to correct for that by assuming that consumption in the utility function is scaled down by the growth factor χ^*, i.e. letting Bernoulli utility $U(\tilde{c})$ in assumption 3 operate on $\tilde{c} \equiv ce^{-\chi^* t}$.

to be equal constant and equal to χ^*, it must then hold that the sum of the growth rates of w_{jt}, h_{jt} and ℓ_{jt} is equal to χ^* i.e.

$$\frac{\dot{w}_{jt}}{w_{jt}} + \frac{\dot{h}_{jt}}{h_{jt}} + \frac{\dot{\ell}_{jt}}{\ell_{jt}} = \chi^*.$$

However, attention to labor has an upper limit $\ell_{jt} \leq 1$, so it cannot grow at a constant rate. Similarly, since learning higher-indexed skills becomes increasingly difficult ($\frac{\partial}{\partial j} g_j < 0$), there exists an index j' above which rate of human capital accumulation will fall below χ^*. Hence, on the balanced growth path $\frac{\dot{w}_{jt}}{w_{jt}} = \chi^*$. When wages for the skills grow at a constant rate, so will the price of the task if done by labor p^L_{jt}, which at some point will surpass the constant price of the task done by machines p^*_j. At this point automation occurs and the task is done by machines. Hence, by (27) all tasks will eventually be automated.

Furthermore, following a similar line of argument, we have faster adoption rates of automated technology than new tasks leading to routine-biased technological change. Indeed, since learning higher-indexed skills becomes increasingly difficult ($\frac{\partial}{\partial j} g_j < 0$), there exists an index above which the rate of human capital accumulation at the technology frontier $h_{N^*_t}$ will be less than that of physical capital X_t. Hence, since $\ell_{N^*_t} \leq 1$, by (31) we have faster adoption rates of automated technology than new tasks leading to routine-biased technological change [5]:

$$\mathbb{E}\left\{\frac{d\widetilde{I}_t}{\widetilde{I}_t}\right\} > \mathbb{E}\left\{\frac{d\widetilde{N}_t}{\widetilde{N}_t}\right\},$$

and thereby

$$\mathbb{E}\left\{\frac{dI^*_t}{I^*_t}\right\} > \mathbb{E}\left\{\frac{dN^*_t}{N^*_t}\right\}.$$

Hence, total automation and labor immiseration follows

$$\operatorname*{plim}_{t\to\infty} n^*_t = 0.$$

We summarize the results in the proposition below.

Proposition 3.7. *Under assumptions 1 to 4 labor immiseration is inevitable, i.e.* $\operatorname*{plim}_{t\to\infty} n^*_t = 0$.
Proof. See the proof by contradiction elaborated above. □

Observe that this labor immiseration scenario is consistent with empirical evidence on effects of automation and returns to education. A crucial assumption is that tasks become increasingly difficult to master ($\frac{\partial}{\partial j} g_j < 0$). This result run contrary to the insight in Acemoglu and Restrepo (2018*e*) who conclude that full automation can be avoided as long as arrival rates of automation technology I_t and new tasks N_t are equal. The reason for this departure is that Acemoglu and Restrepo (2018*e*) assume capital is not too abundant relative to labor. Here, however, as the supply of human-capital-augmented labor is endogenized, we see that as new tasks become increasingly difficult to master, capital becomes more abundant over time, lower price of tasks being performed by machines. Eventually economy becomes fully automated and labor immiserates.

Nevertheless, if learning new skills can become easier over time, we could avoid labor immiseration. Indeed taxing capital and investing the revenue in increasing the productivity of education could be remedy for labor immiseration as detailed in Galor and Moav (2006). Assume for instance that the learning function $g_j(a)$ is augmented by state-enhanced productivity investment $A(T_tX_t)$ financed by a taxation of capital T_tX_t, where $0 < T_t \leq 1$ is the taxation rate and $A'(x) > 0$. By taxing capital, its rate of accumulation is reduced, and by its subsequent investment into education, the corresponding rate for human capital stock is increased. More specifically, for a balanced growth path to exist, the state-enhanced productivity $A(T_tX_t)$ also needs to consider the technology frontier N_t^* where the diminishing marginal productivity of learning crucially must be offset. For instance, for the state-enhanced productivity of the form

$$A(T_tX_t) = e^{\gamma_a T_t X_t}, \qquad \gamma_a > 0,$$

and comparing to the learning function in Assumption 3, we can see that a balanced growth path exists, if $\gamma_a T_t X_t = \gamma_g N_t^*$, i.e. if

$$T_t = \frac{\gamma_g N_t^*}{\gamma_a X_t}.$$

In such a scenario marginal learning productivity is no longer diminishing just as in Grossman et al. (2020) who also find balanced growth paths under such circumstances. Grossman et al. (2020) do not differentiate human capital stock into different skills, however.

However, a hallmark of this balanced growth path is perpetuating job market polarization [6] (cf. e.g. Autor, Katz and Kearney, 2006; Goos, Manning and Salomons, 2014). In such a scenario workers fall into two categories those who have more stock in non-routine skills (high-indexed workers) and those adept at more routine tasks (low-indexed workers). Hence unlike the results in Galor and Moav (2006), which predict a demise of the class structure, the class divide between these to types is perpetuated and intensifies over time. This result is in line with increasing income inequality due to the disproportionate rise of professionals' and managers' compensation in recent decades (cf. e.g. Atkinson, Piketty and Saez, 2011). I detail the result in the next section below.

3.4 Job Market Polarization

Now we use the assumptions we made about the individuals' aggregate stock of human capital $\boldsymbol{h}_{it}$ in (37) and (38), their skill density $f_{h_{it}}(j) \sim e^{-\gamma_{h_i} j}$ and the population's distribution in the endowment parameter $1/\gamma_{h_i} \sim \mathrm{Exp}(\varrho)$. We also assume the following on the technology structure to make the illustration of the rest of analysis easier.

Assumption 5.

$$\vartheta_L(j) = e^{\gamma_L j} \text{ and } \vartheta_M(j) = e^{\gamma_M j} \text{ where } \gamma_L > \gamma_M, \text{ and moreover } F_{\Delta_I} \sim e^{\delta_I j}.$$

We can then prove the following proposition which predicts polarization in the labor market following endowment profiles in human capital.

Proposition 3.8. *Assumptions 3 to 5 hold and workers are cross-productive with distance* $\varsigma > 0$ *such that* $N_t^* + \varsigma > N_t + dN_t$*. Moreover, assume* $r_t > \lambda \mathcal{P}_{\Delta_{N_t^*}} - \sigma^2 > 0$ *such that* $r_t = \lambda \mathcal{P}_{\Delta_{N_t^*}} - \sigma^2 + \epsilon$ *where* $\epsilon > 0$*. Then when* $N_t > \underline{N}$ *for some* $\underline{N} > 0$*, there exists a threshold* $\widetilde{\gamma}_{h_i}$ *in the population given by*

$$\widetilde{\gamma}_{h_i} = \delta_I + \left(\frac{\varepsilon - 1}{\varepsilon}\right)\gamma_L + \gamma_g, \tag{51}$$

such that low-indexed workers with $\gamma_{h_i} \geq \widetilde{\gamma}_{h_i}$ *will have decreasing learning investment* a_{ijt} *along the task index* j *while for high-indexed workers with higher endowment parameter* $\gamma_{h_i} \leq \widetilde{\gamma}_{h_i}$ *learning investments are increasing.*

Proof. See the appendix. □

Figure 7 illustrates the human capital investment profiles of these worker types. There will be within-type heterogeneity [2] leading to differentiated distributional effects of technology adoption within and without groups. As in the static model, agents sort into education based on realized returns and there exists selection bias and sorting gains in schooling. Observe also that low-index agents are more prone to having to change the index interval upon which they invest in human capital. Hence, they are forced to be more occupationally mobile, and penalized for it due to their skills at the automation threshold becoming obsolete. We also see that agents which have comparatively higher endowment in skills that are more difficult to acquire, tend to be more prone to work with new complex technology [3].

Figure 7 goes here.

Since the distribution of the endowment parameter is given by $1/\gamma_{h_i} \sim \text{Exp}(\varrho)$, there are more of low-indexed agents than high indexed ones and thereby the scarcity premium in (33) is larger than one, making the skill premium ambiguous. Observe that Assumption 5 is not crucial to the analysis above, but provides ease in the characterization of the threshold $\widetilde{\gamma}_{h_i}$. The resulting division into high- and low-index types of agents speaks to job-market polarization [6]. Along a balanced growth path where labor income Y_t^L grows at constant rate requires human capital stocks h_{jt} grow at a constant rate. In section 3.3 above we detailed that a taxation of capital and investments in education can accommodate that. However, given the human capital profiles in Proposition (3.8) workers are trapped in their types, i.e. either low or high index. Hence, as physical capital is accumulated the discrepancies between groups will be perpetuated. Moreover, given local cross-productivity, they extend along the analytic continuation of a_{ijt} as I_t^* and N_t^* progress. This continuation for the low-index is decreasing in j and increasing for the high-index, why over time their differences exacerbate. Indeed as $t \to \infty$, the share of income held by the low-index goes towards zero. These results depart from Galor and Moav (2006) who predict the demise of the class structure following common interest by employees and employers in seeing a broadly educated public. In the framework here some high-index workers are better at keeping up with the technological frontier. Hence, the job polarization and a new class divide emerges. This new divide is between high and low-index workers, unlike the old divide between employees and employers.

4 Discussion of Assumptions and Results

It is useful to discuss the relation some of the main assumptions to the results. The consistency of the model with the empirical evidence on returns to education and skill profiles (II) in the introduction section 1 follows directly from assumptions grounded in the psychometric literature on skill formation (I). This consistency is important so that the incentive structure prompting workers to pursue education is empirically sound in the model.

Assumption 1 – i.e. the fact that tasks are gross complements ($0 < \varepsilon < 1$) – together with the comparative advantage of labor in higher-indexed (non-routine) skills, makes the productivity premium less than one. Thus, whether wages of higher-indexed tasks are higher than lower-indexed ones depends on the relative scarcity of their human-capital-augmented labor supply. RBTC [5] follows from ordering tasks from most to least routine and assuming that labor has comparative advantage in non-routine tasks over machines. This setting borrows directly from Autor, Levy and Murnane (2003) and Acemoglu and Restrepo (2018e). Assumption 1 is thus conceptually motivated. Tasks are distinguished on a level of granularity where they complement each other and then are ordered based on their degree of routineness. As such, this assumption guarantees that the whole production of the economy is not performed by only one cheapest task. Moreover, it also guarantees that quality-adjusted investment goods machines, are cheaper, in line with empirical evidence [7]. Allowing tasks being gross substitutes ($\varepsilon > 1$) would, in fact, entail that the skill threshold polarization in the labor market $\widetilde{\gamma}_{h_i}$ in (51) is always positive, guaranteeing the discussed class divide. Assumption 2 functions as a simplification on workers' financial investment decisions and reflect a no arbitrage condition in the capital market.

The learning structure in Assumption 3 is key to the result on falling labor share and inevitable labor immiseration in absence of taxation on physical capital. As tasks become less routine and thus more difficult to master, human capital accumulates at an increasingly slower rate. Consequently, it cannot compete with the lower prices offered by machines in performing routine tasks. It is then no surprise that an improvement in learning efficiency financed through taxing physical capital can alleviate this phenomenon. Assumption 3 creates the key difference of this framework, relative to the one developed in Acemoglu and Restrepo (2018e). In their framework, they assume that capital and labor are in a relative balance in terms of their abundance (see assumption 3 in their paper). In the setting developed here, by viewing labor as human-capital-augmented, we see that workers' endogenous learning choices cannot guarantee such balance. This paper hence stresses the pivotal role of human capital accumulation and human-capital promoting institutions in determining the growth rate similar to the work of Galor and Moav (2004) and Galor, Moav and Vollrath (2009).

Another deviation of the current study from the framework of Acemoglu and Restrepo (2018*e*) is the exogenous arrival rate of automation technology I_t and new tasks N_t. Acemoglu and Restrepo (2018*e*) endogenize the rates of arrival by assuming a pool of scientists being allocated between research and development on automation and new-task creation as illustrated in Acemoglu and Restrepo (2018*e*). As the framework here has introduced human capital, it would be a conceptual violation to assume that the skills of scientists renders them of a different ilk. Nevertheless, Acemoglu and Restrepo (2018*e*) show that for balanced growth paths, equal arrival rate of both technologies is a necessary condition. Hence, by assuming an exogenous yet equal rate of arrival, we abstract away an immaterial mechanism while rigging the growth path in favor of being balanced. However, unbalanced growth emerges despite these favorable conditions.

The volatility structure posited in Assumption 4 is a qualified geometric diffusion. Just as the plain geometric diffusion process, however, the volatility in this setting approaches zero as financial capital stock diminishes indefinitely, which in turn guaranties non-negative values on assets. The particular qualifications here are in place mainly so as to achieve tractability, but nevertheless do not entail any unusual consequences. As illustrated in the proof of the appendix the specific structure is motivated by the solutions found under the assumptions of no asset volatility, and hence has no major bearing on the key results on labor immiseration and job market polarization.

Assumption 5 is merely a functional form specification of the comparative advantage of labor in non-routine tasks. It is similar to assumption $1''$ in Acemoglu and Restrepo (2018*e*). Job market polarization [6] follows mainly from labor's comparative advantage in non-routine tasks and the cross-productivity of skills [a_3]. Crossproductivity articulated in Definition 3.3 plays a central role. Importantly cross-productivity entails that the stock-density of human capital $\boldsymbol{h}_{it}$ extend along its analytic continuation. This continuation for the low-index worker is decreasing in j. In other words, they have more stock in routine stocks. Similarly, the continuation of the stock-density is increasing for the high-index why they have more stock in non-routine skills. As such, over time the difference between the groups exacerbate as routine tasks automate and new non-routine tasks become part of the production process. These results depart from Galor and Moav (2006) who predict the demise of the class structure following common interest by employees and employers in seeing a broadly educated public. In the framework here some high-index workers are better at keeping up with the technological frontier. Hence, the job polarization and a new class divide emerges. This new divide is between high and low-index workers, unlike the old divide between employees and employers.

It is also prudent to relate the results of this paper to a number of other theoretical work – other than Acemoglu and Restrepo (2018e,c) – which attempt to investigate the impact of automation on growth and inequality. I will relate my results here to three other work, namely, Aghion, Jones and Jones (2019), Prettner and Strulik (2020), Grossman et al. (2020). Aghion, Jones and Jones (2019) investigate the impact of AI on growth. Though they do not include human capital or any mechanism for its accumulation, their results are relevant as they consider how growth is impacted under full automation of a finite and discrete number of tasks. In their framework, the growth rate pertaining to the productivity of physical tasks is augmented by the productivity of cognitive tasks. If there is increasing returns to cognitive tasks, then there will be explosive growth in output, either in finite time or asymptotically. In particular, they find that full automation in finite time leads to explosive growth. That does not necessarily occur in the framework developed here, however, as I do not consider a case with increasing returns-to-scale. Indeed, on the balanced growth paths discussed in Sections 3.3 and 3.4, labor could immiserate either in finite time or asymptotically.

Another related work is Prettner and Strulik (2020). They consider a dichotomous case with high and low-ability individuals. Low-ability individuals do not attain tertiary education as the disutility resulted from such an endeavor is too great. In my framework, the two types of high and low-skill also emerge due to their ability type, though within group heterogeneity is absent in Prettner and Strulik (2020). In their framework automation can adversely impact only the low-ability individuals. Since in my framework there is a risk for full automation, the adverse effects are felt by all in such a scenario, however. Nevertheless, the conclusions of Prettner and Strulik (2020) bears some resemblance to the scenario with job market polarization discussed in Section 3.4.

Finally, the framework of Grossman et al. (2020) is closely related and can be seen as a special case of mine. Their scenario only considers one type of human capital and marginal product of human capital investment is constant. Under such a scenario, they find balanced growth paths which exhibit constant shares of production for capital and labor. These conclusions are confirmed here. Furthermore, my framework extend on these insights, where I show that if the marginal product of human capital investment is diminishing then any balanced growth path involves labor immiseration.

5 Concluding Remarks

This study includes considerations of human capital investment based on the conceptual framework of the psychometric literature on skill formation in a task-based environment. Moreover, the focus of the study is shifted towards institutional adoption of new technologies, rather than their arrival rate. In doing so, the model produces an inevitable labor immiseration scenario consistent with several empirical findings on returns to education and decreasing labor share of production. The main mechanism for many of these results is shown to be the faster accumulation rate of physical vis-à-vis human capital. While capital taxation to fund investments in education can hinder labor immiseration, such policies are shown to perpetuate and exacerbate the polarization between low and high index workers. The reason for this result is shown to be cross-productivity of skills at the technological frontier. Indeed, high-index workers can keep up with new tasks, while the situation of low-index workers become increasingly precarious. Future research should include numerical calibration and matching the erosion of labor share over time.

References

Acemoglu, Daron. 1998. "Why do new technologies complement skills? Directed technical change and wage inequality." *The Quarterly Journal of Economics* 113(4):1055–1089.

Acemoglu, Daron and David Autor. 2011. Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of Labor Economics*. Vol. 4 Elsevier pp. 1043–1171.

Acemoglu, Daron and Pascual Restrepo. 2018*a*. "Artificial Intelligence, Automation and Work." *NBER Working Paper Series* No. 24321.

Acemoglu, Daron and Pascual Restrepo. 2018*b*. "Demographics and Automation." *NBER Working Paper Series* No. 24421.

Acemoglu, Daron and Pascual Restrepo. 2018*c*. "Low-Skill and High-Skill Automation." *Journal of Human Capital* 12(2):204–232.

Acemoglu, Daron and Pascual Restrepo. 2018*d*. "Modeling Automation." *NBER Working Paper Series* No. 24321.

Acemoglu, Daron and Pascual Restrepo. 2018*e*. "The race between machine and man: Implications of technology for growth, factor shares and employment." *American Economic Review* 108(6):1488–1542.

Adermon, Adrian and Magnus Gustavsson. 2015. "Job Polarization and Task-Biased Technological Change: Evidence from Sweden, 1975–2005." *The Scandinavian Journal of Economics* 117(3):878–917.

Aghion, Philippe. 2002. "Schumpeterian growth theory and the dynamics of income inequality." *Econometrica* 70(3):855–882.

Aghion, Philippe, Benjamin F Jones and Charles I Jones. 2019. Artificial Intelligence and Economic Growth. In *The Economics of Artificial Intelligence: An Agenda*, ed. Ajay Agrawal, Joshua Gans and Avi Goldfarb. University of Chicago Press.

Aghion, Philippe, Ufuk Akcigit and Peter Howitt. 2014. What do we learn from Schumpeterian growth theory? In *Handbook of Economic Growth*. Vol. 2 Elsevier pp. 515–563.

Ahn, SeHyoun, Greg Kaplan, Benjamin Moll, Thomas Winberry and Christian Wolf. 2018. "When inequality matters for macro and macro matters for inequality." *NBER Macroeconomics Annual* 32(1):1–75.

Ali Akbari, Danial and Thomas Fischer. 2020. "Shaping Inequality: Progressive Taxation under Human Capital Accumulation." *Working paper* .

Almond, Douglas and Janet Currie. 2011. "Killing me softly: The fetal origins hypothesis." *Journal of Economic Perspectives* 25(3):153–72.

Atkinson, Anthony B, Thomas Piketty and Emmanuel Saez. 2011. "Top Incomes in the Long Run of History." *Journal of Economic Literature* 49(1):3–71.

Autor, David. 2013. "The 'Task Approach' to Labor Markets: An Overview." *Journal for Labour Market Research* 46(3):185 – 199.

Autor, David. 2015. "Why Are There Still So Many Jobs? The History and Future of Workplace Automation." *Journal of Economic Perspectives* 29(3):3–30.

Autor, David. 2019. "Work of the Past, Work of the Future." *NBER Working Paper Series* No. 25588.

Autor, David and Anna Salomons. 2018. "Is automation labor-displacing? Productivity growth, employment, and the labor share." *NBER Working Paper Series* No. 24321.

Autor, David and David Dorn. 2013. "The growth of low-skill service jobs and the polarization of the US labor market." *American Economic Review* 103(5):1553–97.

Autor, David, Frank Levy and Richard J Murnane. 2003. "The skill content of recent technological change: An empirical exploration." *The Quarterly Journal of Economics* 118(4):1279–1333.

Autor, David, Lawrence F Katz and Melissa S Kearney. 2006. "The polarization of the US labor market." *American Economic Review* 96(2):189–194.

Baumol, William J. 1967. "Macroeconomics of Unbalanced Growth: the Anatomy of Urban Crisis." *American Economic Review* 57(3):415–426.

Becker, Gary S. 1985. "Human capital, effort, and the sexual division of labor." *Journal of Labor Economics* 3(1, Part 2):S33–S58.

Becker, Gary S. 2009. *Human capital: A theoretical and empirical analysis, with special reference to education.* 3 ed. University of Chicago press. First publication 1964.

Becker, Gary S and Kevin M Murphy. 1992. "The division of labor, coordination costs, and knowledge." *The Quarterly Journal of Economics* 107(4):1137–1160.

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy* 75(4, Part 1):352–365.

Berg, Andrew, Edward F Buffie and Luis Felipe Zanna. 2018. "Should We Fear the Robot Revolution? (The Correct Answer is Yes)." *IMF Working Papers* WP/18/116.

Blanchard, Olivier J. 1985. "Debt, deficits, and finite horizons." *Journal of Political Economy* 93(2):223–247.

Brynjolfsson, Erik and Lorin M Hitt. 2003. "Computing productivity: Firm-level evidence." *Review of Economics and Statistics* 85(4):793–808.

Calvo, Guillermo A and Maurice Obstfeld. 1988. "Optimal time-consistent fiscal policy with finite lifetimes." *Econometrica* 56(2):411–432.

Card, David. 1999. The Causal Effect of Education on Earnings. In *Handbook of Labor Economics.* Vol. 3 Elsevier pp. 1801–1863.

Card, David. 2001. "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems." *Econometrica* 69(5):1127–1160.

Carneiro, Pedro, James J Heckman and Edward J Vytlacil. 2011. "Estimating marginal returns to education." *American Economic Review* 101(6):2754–81.

Caselli, Francesco and Wilbur John Coleman. 2006. "The world technology frontier." *American Economic Review* 96(3):499–522.

Chabé-Ferret, Sylvain. 2015. "Analysis of the bias of matching and difference-in-difference under alternative earnings and selection processes." *Journal of Econometrics* 185(1):110–123.

Comin, Diego and Bart Hobijn. 2004. "Cross-country technology adoption: making the theories face the facts." *Journal of Monetary Economics* 51(1):39–83.

Cortes, G. M., N. Jaimovich and Henry E. Siu. 2017. "Disappearing routine jobs: Who, how, and why?" *Journal of Monetary Economics* 91:69–87. 69.

Cunha, Flavio and James J Heckman. 2007. "The technology of skill formation." *American Economic Review Papers & Proceedings* 97(2):31–47.

Cunha, Flavio and James J Heckman. 2008. "Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation." *Journal of Human Resources* 43(4):738–782.

Cunha, Flavio, James J Heckman, Lance Lochner and Dimitriy V Masterov. 2006. "Interpreting the Evidence on Life Cycle Skill Formation." *Handbook of the Economics of Education* 1:697–812.

Cunha, Flavio, James J Heckman and Susanne M Schennach. 2010. "Estimating the technology of cognitive and noncognitive skill formation." *Econometrica* 78(3):883–931.

Dao, Mai Chi, Ms Mitali Das, Zsoka Koczan and Weicheng Lian. 2017. "Why is labor receiving a smaller share of global income? Theory and empirical evidence." *IMF Working Paper Series* WP/17/169.

David, Paul A. 1990. "The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox." *American Economic Review: Papers & Proceedings* 80(2):355–361.

Dinopoulos, Elias, Constantinos Syropoulos, Bin Xu and Yoto V Yotov. 2011. “Intraindustry trade and the skill premium: Theory and evidence.” *Journal of International Economics* 84(1):15–25.

Elsby, Michael WL, Bart Hobijn and Ayşegül Şahin. 2013. “The decline of the US labor share.” *Brookings Papers on Economic Activity* 2013(2):1–63.

Feng, Andy and Georg Graetz. 2019. “Training Requirements, Automation, and Job Polarization,.” *Economic Journal* Forthcoming.

Foote, C. L. and R. W. Ryan. 2015. “Labor-market polarization over the business cycle.” *NBER Macroeconomics Annual* 29(1):371–413.

Foster, Andrew D and Mark R Rosenzweig. 2010. “Microeconomics of technology adoption.” *Annu. Rev. Econ.* 2(1):395–424.

Galor, Oded and Omer Moav. 2004. “From physical to human capital accumulation: Inequality and the process of development.” *The Review of Economic Studies* 71(4):1001–1026.

Galor, Oded and Omer Moav. 2006. “Das human-kapital: A theory of the demise of the class structure.” *The Review of Economic Studies* 73(1):85–117.

Galor, Oded, Omer Moav and Dietrich Vollrath. 2009. “Inequality in landownership, the emergence of human-capital promoting institutions, and the great divergence.” *The Review of Economic Studies* 76(1):143–179.

Ginther, Donna K. 2000. “Alternative estimates of the effect of schooling on earnings.” *Review of Economics and Statistics* 82(1):103–116.

Goldin, Claudia and Lawrence F Katz. 1998. “The origins of technology-skill complementarity.” *The Quarterly Journal of Economics* 113(3):693–732.

Goldin, Claudia and Lawrence F Katz. 2009. *The race between education and technology*. Harvard University Press.

Goos, Maarten and Alan Manning. 2007. “Lousy and Lovely Jobs: The Rising Polarization of Work in Britain.” *The Review of Economics and Statistics* 89(1):118–133.

Goos, Maarten, Alan Manning and Anna Salomons. 2009. “Job polarization in Europe.” *American Economic Review* 99(2):58–63.

Goos, Maarten, Alan Manning and Anna Salomons. 2014. “Explaining job polarization: Routine-biased technological change and offshoring.” *American Economic Review* 104(8):2509–26.

Graetz, Georg and Guy Michaels. 2018. “Robots at work.” *Review of Economics and Statistics* 100(5):753–768.

Graff Zivin, Joshua and Matthew Neidell. 2013. “Environment, health, and human capital.” *Journal of Economic Literature* 51(3):689–730.

Grossman, Gene M, Elhanan Helpman, Ezra Oberfield and Thomas Sampson. 2020. "Endogenous Education and Long-Run Factor Shares." *NBER Working Paper Series* No. 27031.

Guvenen, Fatih, Burhanettin Kuruscu and Serdar Ozkan. 2014. "Taxation of human capital and wage inequality: A cross-country analysis." *Review of Economic Studies* 81(2):818–850.

Harrigan, James, Ariell Reshef and Farid Toubal. 2016. "The March of the Techies: Technology, Trade, and Job Polarization in France, 1994-2007." *NBER Working Paper Series* No. 18334.

Heckman, James J. 1976. "A Life-Cycle Model of Earnings, Learning, and Consumption." *Journal of Political Economy* 84(4, Part 2):S9–S44.

Heckman, James J, Daniel Schmierer and Sergio Urzua. 2010. "Testing the correlated random coefficient model." *Journal of Econometrics* 158(2):177–203.

Heckman, James J, John Eric Humphries and Gregory Veramendi. 2018. "Returns to education: The causal effects of education on earnings, health, and smoking." *Journal of Political Economy* 126(S1):S197–S246.

Heckman, James J, Lance J Lochner and Petra E Todd. 2006. "Earnings functions, rates of return and treatment effects: The Mincer equation and beyond." *Handbook of the Economics of Education* 1:307–458.

Heckman, James J, Lance J Lochner and Petra E Todd. 2008. "Earnings functions and rates of return." *Journal of Human Capital* 2(1):1–31.

Heckman, James J and Richard Robb Jr. 1985. "Alternative methods for evaluating the impact of interventions: An overview." *Journal of Econometrics* 30(1-2):239–267.

Helmers, Christian and Manasa Patnam. 2011. "The formation and evolution of childhood skill acquisition: Evidence from India." *Journal of Development Economics* 95(2):252–266.

Hershbein, Brad and Lisa B Kahn. 2018. "Do recessions accelerate routine-biased technological change? Evidence from vacancy postings." *American Economic Review* 108(7):1737–72.

Huggett, Mark, Gustavo Ventura and Amir Yaron. 2006. "Human capital and earnings distribution dynamics." *Journal of Monetary Economics* 53(2):265–290.

Huggett, Mark, Gustavo Ventura and Amir Yaron. 2011. "Sources of lifetime inequality." *American Economic Review* 101(7):2923–54.

Itskhoki, Oleg and Benjamin Moll. 2019. "Optimal development policies with financial frictions." *Econometrica* 87(1):139–173.

Kalemli-Ozcan, Sebnem, Harl E Ryder and David N Weil. 2000. "Mortality decline, human capital investment, and economic growth." *Journal of Development Economics* 62(1):1–23.

Kaplan, Greg, Benjamin Moll and Giovanni L Violante. 2018. "Monetary policy according to HANK." *American Economic Review* 108(3):697–743.

Karabarbounis, Loukas and Brent Neiman. 2014. "The Global Decline of the Labor Share." *Quarterly Journal of Economics* 129(1):61–103.

Katz, Lawrence F and David Autor. 1999. "Changes in the wage structure and earnings inequality." *Handbook of Labor Economics* 3:1463–1555.

Krebs, Tom, Moritz Kuhn and Mark LJ Wright. 2015. "Human capital risk, contract enforcement, and the macroeconomy." *American Economic Review* 105(11):3223–72.

Krusell, Per, Lee E Ohanian, José-Víctor Ríos-Rull and Giovanni L Violante. 2000. "Capital-skill complementarity and inequality: A macroeconomic analysis." *Econometrica* 68(5):1029–1053.

León-Ledesma, Miguel A, Peter McAdam and Alpo Willman. 2010. "Identifying the elasticity of substitution with biased technical change." *American Economic Review* 100(4):1330–57.

Lindahl, Mikael, Mårten Palme, Sofia Sandgren-Massih and Anna Sjögren. 2014. "A test of the Becker-Tomes model of human capital transmission using microdata on four generations." *Journal of Human Capital* 8(1):80–96.

Lindahl, Mikael, Mårten Palme, Sofia Sandgren-Massih and Anna Sjögren. 2015. "Long-term intergenerational persistence of human capital an empirical analysis of four generations." *Journal of Human Resources* 50(1):1–33.

Long, Jason and Joseph Ferrie. 2007. "The path to convergence: intergenerational occupational mobility in Britain and the US in three eras." *The Economic Journal* 117(519):C61–C71.

Long, Jason and Joseph Ferrie. 2013. "Intergenerational occupational mobility in Great Britain and the United States since 1850." *American Economic Review* 103(4):1109–37.

Lucas Jr, Robert E and Benjamin Moll. 2014. "Knowledge growth and the allocation of time." *Journal of Political Economy* 122(1):1–51.

Ludwig, Alexander, Thomas Schelkle and Edgar Vogel. 2012. "Demographic change, human capital and welfare." *Review of Economic Dynamics* 15(1):94–107.

Madsen, Jakob B. 2014. "Human capital and the world technology frontier." *Review of Economics and Statistics* 96(4):676–692.

Mincer, Jacob. 1958. "Investment in human capital and personal income distribution." *Journal of Political Economy* 66(4):281–302.

Mincer, Jacob. 1974. "Schooling, Experience, and Earnings." *Human Behavior & Social Institutions* (No. 2).

Nuño, Galo and Benjamin Moll. 2018. "Social optima in economies with heterogeneous agents." *Review of Economic Dynamics* 28:150–180.

OECD. 2015. "The Labour Share in G20 Economies." *Report prepared for the G20 Employment Working Group* .

Piketty, Thomas. 2014. *Capital in the 21st Century*. Harvard University Press Cambridge, MA.

Prettner, Klaus and Holger Strulik. 2020. "Innovation, automation, and inequality: Policy challenges in the race against the machine." *Journal of Monetary Economics* 116:249–265.

Psacharopoulos, George and Harry Antony Patrinos. 2004. "Returns to investment in education: a further update." *Education Economics* 12(2):111–134.

Psacharopoulos, George and Harry Antony Patrinos. 2018. "Returns to investment in education: a decennial review of the global literature." *Education Economics* pp. 1–14.

Rosen, Sherwin. 1976. "A Theory of Life Earning." *Journal of Political Economy* 84(4, Part 2):S45–S67.

Rosen, Sherwin. 1983. "Specialization and Human Capital." *Journal of Labor Economics* 1(1):43–49.

Schwellnus, Cyrille, Mathilde Pak, Pierre-Alain Pionnier and Elena Crivellaro. 2018. "Labour share developments over the past two decades." *OECD Working Papers* No. 1503.

Stokey, Nancy L. 1988. "Learning by Doing and the Introduction of New Goods." *Journal of Political Economy* 96(4):701–717.

Stokey, Nancy L. 2018. "Technology, Skill, and the Wage Structure." *Journal of Human Capital* 12(2):343–384.

Susskind, Daniel. 2017. "A model of technological unemployment." *University of Oxford, Department of Economics Discussion Paper,* Accessed 2019-02-17 at https://www.economics.ox.ac.uk/materials/papers/15126/819-susskind-a-model-of-technological-unemploymentjuly-2017.pdf.

Turner, Chad, Robert Tamura, Curtis J Simon and Sean Mulholland. 2018. "Dynastic Human Capital and Black-White Earnings Differentials in the United States, 1940–2000." *Journal of Human Capital* 12(2):385–430.

Wallenius, Johanna. 2011. "Human capital accumulation and the intertemporal elasticity of substitution of labor: How large is the bias?" *Review of Economic Dynamics* 14(4):577–591.

Willis, Robert J. 1986. "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions." *Handbook of Labor Economics* 1:525–602.

Yaari, Menahem E. 1965. "Uncertain lifetime, life insurance, and the theory of the consumer." *The Review of Economic Studies* 32(2):137–150.

Zeira, Joseph. 1998. "Workers, machines, and economic growth." *The Quarterly Journal of Economics* 113(4):1091–1117.

A Proofs

The mathematics are combinations of techniques partly in Stokey (1988, 2018) and partly in Lucas Jr and Moll (2014) and Nuño and Moll (2018). Considering a continuum of types in ”knowledge capital” among workers in Stokey (1988) is very similar to the approach that I am proposing here. Stokey (1988) considers a representative-agent partial-equilibrium deterministic learning-by-doing scheme. In contrast, Stokey (2018) develops a deterministic general equilibrium structure yet without any mechanism for learning. I develop a heterogeneous-agent stochastic general-equilibrium approach with skill acquisition by borrowing mean-field techniques employed in Lucas Jr and Moll (2014) and Nuño and Moll (2018). Lucas Jr and Moll (2014) provides a dynamic general equilibrium learning-by-imitation scheme knowledge accumulation with drift-jump stochasticity. However, their model considers only a single measure of productivity as proxy for human capital stock and thus does not deal with plurality of skills which is essential in dealing with risks of devaluation. Moreover, in their model stochasticity is a result of the meetings with higher productivity individuals and hence they are not facing and possibility of their skills becoming obsolete. Nuño and Moll (2018) provides a general setting for how to carry out aggregation in environments with *ex-post* idiosyncratic wealth risk.

A.1 Proofs on the Expository Static Model

We can deduce the following lemma which states that not all of attention to a particular skill will go exclusively towards learning that skill. In other words, if attention is devoted to a particular skill ($p_j \neq 0$), then some attention is given to labor with that skill ($\ell_j \neq 0$).

Lemma A.1. *Let a_j^* be the attention to learning in skill j that maximizes expected utility in (4). Then it follows that:*

$$0 \leq a_j^* < p_j. \tag{52}$$

Proof of Lemma A.1. We know that $c = y_1 + y_2$ and

$$y_1 = w_1 h_1 (p - a_1), \qquad\qquad y_2 = w_2 h_2 (1 - p - a_2).$$

Then by 2 we have that

$$\frac{\partial}{\partial a_j} y_j(a_j) = w_j h_{j,0} \left(g_j'(a_j)(p_j - a_j) - g_j(a_j) \right), \tag{53}$$

where $p_1 = p$ and $p_2 = 1 - p$. Therefore, $\frac{\partial}{\partial a_j} y_j(p_j) < 0$. Moreover,

$$\frac{\partial^2}{\partial a_j^2} y_j(a_j) = w_j h_{j,0} (g_j''(a_j)(p_j - a_j) - 2 g_j'(a_j)) < 0.$$

By $\frac{\partial}{\partial a_j} y_j(p_j) < 0$ and $\frac{\partial^2}{\partial a_j^2} y_j(a_j) < 0$ we get (52). □

Using this lemma we then can prove the following proposition.

Proposition A.1. *If job description p is fixed, the allocation of time between learning a_j and labor $\ell_j = p_j - a_j$ is independent of wage rates w_j, human capital endowments $h_{j,0}$ and obsolescence probabilities ξ_j of each skill, and only depends on the characteristics of the learning functions g_j, $j = 1, 2$.*

Proof of Proposition A.1. Since job description is fixed expected utility is maximized when income streams from the skills $y_j, j = 1, 2$ are maximized. Hence, maximization is independent of the obsolescence probabilities. By (53), we see that $w_j h_{j,0}$ is immaterial to the optimization and acts as a scaling constant. If there is an interior solution it is given by,

$$\frac{\partial}{\partial a_j}\mathbb{E}(U(c)) = 0 \Rightarrow \frac{\partial}{\partial a_j} y_j = 0$$

which yields

$$\frac{g_j'(a_j^*)}{g_j(a_j^*)} = \frac{1}{p_j - a_j^*} \tag{54}$$

and thus the choice of attention to learning a_j^* only depends on the learning function g_j. Otherwise, by Lemma A.1, $a_j^* = 0$. Thus, the proposition is proven. □

The proposition above in essence states that wage rates, human capital endowments and obsolescence probabilities enter workers' decision making if and only if they can freely decide how much their attention is devoted to the routine (bookkeeping) and non-routine (software development) tasks respectively. If she is forced to divide the attention according to some exogenous scheme, then her labor and learning decisions only by the learning structure of each skill.

Proof of Proposition 3.2. In order to prove this result we resolve the problem (4) with the secondary condition 1. The Lagrangian is given by:

$$\mathcal{L} = \mathbb{E}(U(c)) - \mu\,(a_1 + \ell_1 + a_2 + \ell_2 - 1)$$

where μ is the Lagrangian multiplier. First-order conditions are given by:

$$\frac{\partial \mathcal{L}}{\partial x_j} = 0 \Rightarrow (1-\xi_j)\frac{\partial y_j}{\partial x_j}\left[U'(y_1+y_2)(1-\xi_j) + U'(y_j)\xi_j\right] = \mu \text{ where } x_j \in \{a_j, \ell_j\}. \tag{55}$$

where

$$\text{(a)}\ \frac{\partial y_j}{\partial a_j} = w_j g_j'(a_j) h_{j,0}\ell_j \qquad \text{(b)}\ \frac{\partial y_j}{\partial \ell_j} = w_j g_j(a_j) h_{j,0} \tag{56}$$

Using (55) for a_j and ℓ_j we get,

$$\ell_j^* = \frac{g_j(a_j^*)}{g_j'(a_j^*)}. \tag{57}$$

Differentiating (57) with respect to $h_{1,0}$ yields

$$\frac{\partial \ell_j^*}{\partial h_{1,0}} = \frac{(g_j'(a_j^*))^2 - g_j(a_j^*)g_j''(a_j^*)}{(g_j'(a_j^*))^2} \cdot \frac{\partial a_j^*}{\partial h_{1,0}} \tag{58}$$

which imply

$$\text{sign}\left(\frac{\partial \ell_j^*}{\partial h_{1,0}}\right) = \text{sign}\left(\frac{\partial a_j^*}{\partial h_{1,0}}\right) \tag{59}$$

since $g'' < 0$. Differentiating (1) at the optimum with respect to $h_{1,0}$ we get:

$$\frac{\partial}{\partial h_{1,0}}(a_1^* + \ell_1^*) = -\frac{\partial}{\partial h_{1,0}}(a_2^* + \ell_2^*) \tag{60}$$

where $a_j^* + \ell_j^* = p_j^*$. By (58) and (60) we get

$$\text{sign}\left(\frac{\partial a_1^*}{\partial h_{1,0}}\right) = -\,\text{sign}\left(\frac{\partial a_2^*}{\partial h_{1,0}}\right). \tag{61}$$

Finally, we differentiate (55) for a_1 and ℓ_1 with respect to $h_{1,0}$ and setting the left-hand sides equal to each other and inserting (56) we get the following

$$\frac{\partial}{\partial h_{1,0}}(a_1^* + \ell_1^*) = -\frac{w_1 g_1(a_1^*)\ell_1^* \overbrace{[U''(y_1+y_2)(w_1h_1 - w_2h_2)(1-\xi_1)(1-\xi_2) + (1-\xi_1)\xi_2 U''(y_1)]}^{\equiv Q}}{U''(y_1+y_2)(w_1h_1-w_2h_2)^2(1-\xi_1)(1-\xi_2) + (1-\xi_2)\xi_1 U''(y_2)(w_2h_2)^2}.$$

Since $U'' < 0$, $\frac{\partial}{\partial h_{1,0}}(a_1^* + \ell_1^*) > 0$ if and only if $Q > 0$, which yields the condition (7). Then, (6) follows by (59), (60) and (61). Equalities in (6) holds when the equality in (7) holds. □

Proof of Proposition 3.3. The proof mainly follows from the first-order condition (5). First, we consider the condition at $\underline{h}_{1,0}$ which yields,

$$\frac{w_1 g_1(0)\underline{h}_{1,0}}{w_2 g_2(a_2^*)h_{2,0}} = 1.$$

Since $g_1(0) = 1$ by (2), we have that

$$\frac{\underline{h}_{1,0}}{h_{2,0}} = \frac{w_2}{w_1} \cdot g_2(a_2^*).$$

Once again by (2), we have that $g_2(0) = 1$ and $g_2' > 0$, and hence $g_2(a_2^*) \geq 1$. It follows then that

$$\frac{\underline{h}_{1,0}}{h_{2,0}} \geq \frac{w_2}{w_1}. \tag{62}$$

Similarly the condition (5) at $\bar{h}_{1,0}$ yields

$$\frac{\bar{h}_{1,0}}{h_{2,0}} = \frac{w_2}{w_1} \cdot \frac{1}{g_1(a_1^*)},$$

and hence

$$\frac{\bar{h}_{1,0}}{h_{2,0}} \leq \frac{w_2}{w_1}. \tag{63}$$

Then by (62) and (63) we have that $\bar{h}_{1,0} \leq \underline{h}_{1,0}$. However, by construction $\underline{h}_{1,0} \leq \bar{h}_{1,0}$. Hence, $\underline{h}_{1,0} = \bar{h}_{1,0}$, and the proposition is proven. □

A.2 Proofs on the Full Model

Proof of Proposition 3.4. The profit function for the producer given (21) and (12) is the following,

$$\pi_{jt} = \begin{cases} p_{jt}\vartheta_M(j)k_{jt} - r_{jt}k_{jt} & \text{for } j \in [N_t - 1, I_t^*] \\ p_{jt}\vartheta_L(j)h_{jt}\ell_{jt} - w_{jt}h_{jt}\ell_{jt} & \text{for } j \in [I_t^*, N_t] \end{cases} \tag{64}$$

Perfect competition implies zero profits ($\pi_{jt} = 0$) and hence we have,

$$r_{jt} = p_{jt}\vartheta_M(j) \text{ and } w_{jt} = p_{jt}\vartheta_L(j)$$

for $j \in [N_t - 1, I_t^*]$ and $j \in [I_t^*, N_t]$ respectively. Utilizing Lemma 3.1 we then get

$$r_{jt} = P_t \left(\frac{Y_t}{k_{jt}}\right)^{\frac{1}{\epsilon}} \vartheta_M^{\frac{\varepsilon-1}{\varepsilon}}(j) \text{ and } w_{jt} = P_t \left(\frac{Y_t}{h_{jt}\ell_{jt}}\right)^{\frac{1}{\epsilon}} \vartheta_L^{\frac{\varepsilon-1}{\varepsilon}}(j), \tag{65}$$

for $j \in [N_t - 1, I_t^*]$ and $j \in [I_t^*, N_t]$ respectively. Employing Assumption 2, we realize that

$$\frac{\vartheta_M^{\frac{\varepsilon-1}{\varepsilon}}(j)}{k_{jt}^{\frac{1}{\varepsilon}}} = R_t,$$

where R_t is independent of j. Hence,

$$\frac{\vartheta_M^{\varepsilon-1}(j)}{R_t^{\varepsilon}} = k_{jt}. \tag{66}$$

Integrating both sides over $j \in [N_t - 1, I_t^*]$ - by the capital market clearing condition (22) - and rearranging we get the following expression for R_t

$$R_t = \left(\frac{\int_{N_t-1}^{I_t^*} \vartheta_M^{\varepsilon-1}(j)dj}{X_t}\right)^{\frac{1}{\varepsilon}}. \tag{67}$$

Then (25) and (26) follows from (65), (66) and (67). □

The proof of Proposition 3.6 is performed separately for three scenarios: first no asset nor obsolescence risk, second only obsolescence risk, and third with both asset and obsolescence risk. The first and second scenarios are proper viscosity solutions (i.e. both viscosity sub- and supersolutions), while the the third scenario can only be shown to be a viscosity supersolution. All scenarios are solved for two types of agents: those with diversifiers (i.e. those who learn to perform all tasks pertaining to labor) and specializers (i.e. those who learn to perform a subset of tasks pertaining to labor).

Scenario One - No Asset or Obsolescence Risk: We first derive the optimal decision functions for this agent type under the assumption of no asset and jump risks, and then move onto discuss a solution with a specific risk function. Hence, the following assumption:

Assumption 4′.

No asset or obsolescence risk: $\sigma_i(x_{it}, \boldsymbol{h}_{it}) = \lambda = 0.$

Type 1 - Diversifiers: Diversifiers have manageable human-capital-augmented wage rate $w_{jt}h_{ijt} = w_{j't}h_{ij't}$ on the whole task spectrum performed by labor, i.e. for all $j, j' \in [I_t^*, N_t^*]$.

Proposition A.2. *Under assumptions 3 and 4′, the agent with manageable human-capital-augmented wage rate facing the problem in (34) subject to (36), (39), (16), (17) and (41), is a diversifier and acts in accordance with the following decision rules:*

$$a_{ijt}^* = \frac{1}{\xi}\left(1 - \sqrt{(4e^{\gamma_g j}\xi r_t - 1)\frac{w_{jt}h_{ijt}}{\overline{wh}_{it}}}\right), \qquad \ell_{ijt}^* = \frac{1}{n_t^*} - a_{ijt}^* \tag{68}$$

$$c_{it} = \frac{1}{\alpha}\left(1 - D_{1it} - D_{2t}\widetilde{Q}\right), \quad \textit{where } \widetilde{Q} = x_{it} + \int_{I_t^*}^{N_t^*} D_{ijt}\sqrt{2h_{ijt}}dj \tag{69}$$

and D_{1it}, D_{2t} and D_{ijt} are constants given by

$$D_{ijt} = \sqrt{\frac{2e^{\gamma_g j}\overline{wh}_{it}w_{ijt}}{4\xi r_t - e^{-\gamma_g j}}}, \qquad D_{2t} = \frac{\alpha(\rho - 2r_t)}{6\,(p_t^c)^2} \tag{70}$$

$$D_{1t} = \frac{(\rho - 2r_t)\left[1 + \frac{\alpha\overline{wh}_{it}(1 - \frac{1}{\xi}n_t^*)}{p_t^c}\right]}{p_t^c\,(4\rho - 5r_t)} \tag{71}$$

Proof of Proposition A.2. For ease of notation we suppress the agent and time indexes i and t and assume that the agent takes r_t and w_t as given. Then the general problem (34) is autonomous with the following HJB equation:

$$\rho V = \max_{a_j,\ell_j,c}\Bigg\{U(c) + \frac{\partial V}{\partial x}\left(rx + \int_{I^*}^{N_t} w_j h_j \ell_j dj - p^c c\right) + \int_{I^*}^{N_t}\frac{\partial V}{\partial h_j}g_j(a_j)h_j dj + \frac{1}{2}\sigma_i^2(x_{it}, \boldsymbol{h}_{it})\frac{\partial^2 V}{\partial x^2}$$
$$+\lambda\left[\int_0^{n_t^*}(V(., I^* + s, .) - V)f_{\Delta_I}(s)ds + \int_0^{1-n_t^*}(V(., N + s) - V)f_{\Delta_N}(s)ds\right]\Bigg\} \tag{72}$$

where $V \equiv V(x, \boldsymbol{h}, I^*, N)$ is the optimal current-value function. The Lagrangian for the right-hand side is then given by:

$$\mathcal{L} = U(c) + \frac{\partial V}{\partial x}\left(rx + \int_{I^*}^{N_t} w_j h_j \ell_j dj - p^c c\right) + \int_{I^*}^{N_t} \frac{\partial V}{\partial h_j} g_j(a_j) h_j dj + \frac{1}{2}\sigma_i^2(x_{it}, \boldsymbol{h}_{it})\frac{\partial^2 V}{\partial x^2}$$
$$+ \lambda\left[\int_0^{n_t^*}(V(., I^* + s, .) - V) f_{\Delta_I}(s) ds + \int_0^{1-n_t^*}(V(., N + s) - V) f_{\Delta_N}(s) ds\right]$$
$$- \mu\left(\int_{I_t^*}^{N_t}(a_j + \ell_j) dj - 1\right) \tag{73}$$

where μ is the Lagrangian multiplier. To find the stationary points we differentiate $\mathcal{L}$ with respect to a_j, ℓ_j and c and set the result equal to zero. For differentiation with respect to a_j and ℓ_j we need to employ the Euler-Lagrange equation for the index dimension j since we are optimizing a functional.[6] Thus, we get the following equations,

$$U'(c) = p^c \frac{\partial V}{\partial x} \tag{74}$$
$$\frac{\partial V}{\partial h_j} g_j'(a_j) h_j = \mu \tag{75}$$
$$\frac{\partial V}{\partial x} w_j h_j = \mu \tag{76}$$

derived from $\partial \mathcal{L}/\partial c = 0$, $\partial \mathcal{L}/\partial a_j = 0$ and $\partial \mathcal{L}/\partial \ell_j = 0$ respectively. First-order conditions (75) and (76) need to holds for all $j \in [I_t^*, N_t]$. Observe that since the human-capital-augmented wage rate for this type of agent is manageable ($w_{jt} h_{ijt} = w_{j't} h_{ij't}, \forall j, j' \in [I_t^*, N_t]$), we have that:

$$w_j h_j = \overline{wh} \text{ for all } j \in [I_t^*, N_t] \tag{77}$$

and therefore (76) trivially holds. Then by (75) and (76) we have that

$$\frac{g_j'(a_j)}{g_k'(a_k)} = \frac{w_j}{w_k} \cdot \frac{\frac{\partial V}{\partial h_k}}{\frac{\partial V}{\partial h_j}} \tag{78}$$

for $j \neq k$. By Assumption 3, (74), (75) and (76) yield

$$c^* = \frac{1}{\alpha}\left(1 - p^c \frac{\partial V}{\partial x}\right) \tag{79}$$
$$a_j = \frac{1}{\xi}\left(1 - e^{\gamma_g j} w_j \frac{\frac{\partial V}{\partial x}}{\frac{\partial V}{\partial h_j}}\right) \tag{80}$$

[6]One could indeed optimize using different types of functional derivatives given some perturbation and provide a more granular proof for the differentiation as done by e.g. Lucas Jr and Moll (2014). Nevertheless, the Euler-Lagrange equation is indeed itself proven using directional perturbation derivatives. Hence, the application of said equation is well-motivated even though usually it is formulated along the time dimension.

respectively. Then by (78) and (80) we get

$$a_j = \frac{1}{\xi}\left(1 - \frac{w_j}{w_k} \cdot \frac{\frac{\partial V}{\partial h_k}}{\frac{\partial V}{\partial h_j}}(1-\xi a_k)e^{-\gamma_g(k-j)}\right). \tag{81}$$

Hence the agent's total attention to learning is given by,

$$\int_{I^*}^{N} a_j dj = \frac{1}{\xi}\left(n^* - \frac{e^{-\gamma_g k}}{w_k} \cdot \frac{\partial V}{\partial h_k}(1-\xi a_k)\int_{I^*}^{N} \frac{e^{\gamma_g j}}{\frac{\partial V}{\partial h_j}} dj\right) \tag{82}$$

and total attention to labor - by the budget (41) - is given by

$$\int_{I^*}^{N} \ell_j dj = 1 - \int_{I^*}^{N} a_j dj. \tag{83}$$

Inserting (81) for skill k into (82) we arrive at the following expression for total attention to learning:

$$\int_{I^*}^{N} a_j dj = \frac{1}{\xi}\left(n^* - \frac{\partial V}{\partial x}\int_{I^*}^{N} \frac{e^{\gamma_g j} w_j}{\frac{\partial V}{\partial h_j}} dj\right) \tag{84}$$

which indicates that optimal attention to learning for skill is the following:

$$a_j^* = \frac{1}{\xi}\left(1 - e^{\gamma_g j} w_j \frac{\frac{\partial V}{\partial x}}{\frac{\partial V}{\partial h_j}}\right) \tag{85}$$

and thereby

$$g_j(a_j^*) = \frac{-e^{\gamma_g j}}{2\xi}\left(1 - \left(e^{\gamma_g j} w_j \frac{\frac{\partial V}{\partial x}}{\frac{\partial V}{\partial h_j}}\right)^2\right) \tag{86}$$

Thus we have that

$$\int_{I^*}^{N_t} \frac{\partial V}{\partial h_j} g_j(a_j^*) h_j dj = \frac{1}{2\xi}\int_{I^*}^{N}\left(e^{-\gamma_g j} h_j \frac{\partial V}{\partial h_j} - \overline{wh} w_j e^{\gamma_g j}\frac{\left(\frac{\partial V}{\partial x}\right)^2}{\frac{\partial V}{\partial h_j}}\right) dj. \tag{87}$$

Furthermore, total labor income by (77), (83) and (84) is given by

$$\int_{I^*}^{N} w_j h_j \ell_j^* dj = \overline{wh}\left(1 - \frac{1}{\xi}\left(n^* - \frac{\partial V}{\partial x}\int_{I^*}^{N} \frac{e^{\gamma_g j} w_j}{\frac{\partial V}{\partial h_j}} dj\right)\right). \tag{88}$$

Moreover, by Assumption 3 and (79) we have that

$$U(c^*) = \frac{1}{2\alpha}\left(1 - \left(p^c \frac{\partial V}{\partial x}\right)^2\right) \tag{89}$$

Finally, inserting (79), (87), (88) and (89) into the HJB equation (72) yields then following partial integro-differential equation (PIDE):

$$\rho V = \frac{1}{2\alpha}\left(1-\left(p^c\frac{\partial V}{\partial x}\right)^2\right)+\frac{\partial V}{\partial x}\left[rx+\overline{wh}\left(1-\frac{1}{\xi}\left(n^*-\frac{\partial V}{\partial x}\int_{I^*}^{N}\frac{e^{\gamma_g j}w_j}{\frac{\partial V}{\partial h_j}}dj\right)\right)-\frac{p^c}{\alpha}\left(1-p^c\frac{\partial V}{\partial x}\right)\right]$$
$$+\frac{1}{2\xi}\int_{I^*}^{N}\left(e^{-\gamma_g j}h_j\frac{\partial V}{\partial h_j}-\overline{wh}w_j e^{\gamma_g j}\frac{\left(\frac{\partial V}{\partial x}\right)^2}{\frac{\partial V}{\partial h_j}}\right)dj+\frac{1}{2}\sigma_i^2(x_{it},\boldsymbol{h}_{it})\frac{\partial^2 V}{\partial x^2}$$
$$+\lambda\left[\int_0^{n_t^*}(V(.,I^*+s,.)-V)f_{\Delta_I}(s)ds+\int_0^{1-n_t^*}(V(.,N+s)-V)f_{\Delta_N}(s)ds\right] \tag{90}$$

which by Assumption 4′ is reduced to the following PDE,

$$\rho V = \frac{1}{2\alpha}\left(1-\left(p^c\frac{\partial V}{\partial x}\right)^2\right)+\frac{\partial V}{\partial x}\left[rx+\overline{wh}\left(1-\frac{1}{\xi}\left(n^*-\frac{\partial V}{\partial x}\int_{I^*}^{N}\frac{e^{\gamma_g j}w_j}{\frac{\partial V}{\partial h_j}}dj\right)\right)-\frac{p^c}{\alpha}\left(1-p^c\frac{\partial V}{\partial x}\right)\right]$$
$$+\frac{1}{2\xi}\int_{I^*}^{N}\left(e^{-\gamma_g j}h_j\frac{\partial V}{\partial h_j}-\overline{wh}w_j e^{\gamma_g j}\frac{\left(\frac{\partial V}{\partial x}\right)^2}{\frac{\partial V}{\partial h_j}}\right)dj. \tag{91}$$

To solve this problem we employ the following *ansatz* for the optimal current-value function

$$V(x,\boldsymbol{h};I^*,N)=D_0+D_1\tilde{Q}+D_2\tilde{Q}^2 \tag{92}$$

where $\tilde{Q}$ is described by (69) and (70).[7] We have then

$$\frac{\partial V}{\partial x}=D_1+2D_2\tilde{Q} \text{ and } \frac{\partial V}{\partial h_j}=\frac{D_j}{\sqrt{2h_j}}\cdot\frac{\partial V}{\partial x} \tag{93}$$

for $j\in[I^*,N]$ where D_j is given by (70). Employing (93), we get

$$\frac{1}{2\alpha}\left(1-\left(p^c\frac{\partial V}{\partial x}\right)^2\right)=\frac{1}{2\alpha}(1-D_1^2)+\frac{p^cD_1D_2}{\alpha}\tilde{Q}+\frac{2(p^cD_2)^2}{\alpha}\tilde{Q}^2 \tag{94}$$

$$rx\frac{\partial V}{\partial x}=rD_1x+2rD_2x\tilde{Q} \tag{95}$$

$$\frac{\partial V}{\partial x}\overline{wh}\left(1-\frac{1}{\xi}\left(n^*-\frac{\partial V}{\partial x}\int_{I^*}^{N}\frac{e^{\gamma_g j}w_j}{\frac{\partial V}{\partial h_j}}dj\right)\right)=\overline{wh}(1-\frac{1}{\xi}n^*)D_1+2\overline{wh}(1-\frac{1}{\xi}n^*)D_2\tilde{Q}$$
$$+\frac{\overline{wh}}{\xi}D_1\int_{I^*}^{N}\frac{e^{\gamma_g j}w_j}{D_j}\sqrt{2h_j}dj+2\frac{\overline{wh}}{\xi}D_2\int_{I^*}^{N}\frac{e^{\gamma_g j}w_j}{D_j}\sqrt{2h_j}dj\tilde{Q} \tag{96}$$

$$\frac{\partial V}{\partial x}\frac{p^c}{\alpha}\left(1-p^c\frac{\partial V}{\partial x}\right)=\frac{p^c}{\alpha}D_1-\frac{(p^cD_1)^2}{\alpha}+2\left(\frac{p^c}{\alpha}D_2-\frac{(p^c)^2}{\alpha}D_1D_2\right)\tilde{Q}-\frac{4\,(p^cD_2)^2}{\alpha}\tilde{Q}^2 \tag{97}$$

$$\frac{1}{2\xi}\int_{I^*}^{N}e^{-\gamma_g j}h_j\frac{\partial V}{\partial h_j}dj=\frac{D_1}{2\xi}\int_{I^*}^{N}\frac{e^{-\gamma_g j}D_j}{2}\sqrt{2h_j}dj+\frac{D_2}{\xi}\int_{I^*}^{N}\frac{e^{-\gamma_g j}D_j}{2}\sqrt{2h_j}dj\tilde{Q} \tag{98}$$

$$\frac{1}{2\xi}\int_{I^*}^{N}\overline{wh}w_je^{\gamma_g j}\frac{\left(\frac{\partial V}{\partial x}\right)^2}{\frac{\partial V}{\partial h_j}}dj=\frac{D_1}{2\xi}\int_{I^*}^{N}\frac{\overline{wh}w_je^{\gamma_g j}}{D_j}\sqrt{2h_j}dj+\frac{D_2}{\xi}\int_{I^*}^{N}\frac{\overline{wh}w_je^{\gamma_g j}}{D_j}\sqrt{2h_j}dj\tilde{Q}. \tag{99}$$

[7] Observe that the semi-colon in $V(x,\boldsymbol{h};I^*,N)$ is there to indicate that by Assumption (4′) I^* and N are now just constants.

Observe also that

$$\frac{\overline{wh}e^{\gamma_g j}w_j}{2\xi D_j} + \frac{e^{-\gamma_j}D_j}{4\xi} = rD_j. \tag{100}$$

Inserting (92) and (94)-(99) into the PDE (91) and employing (100) yields:

$$\begin{aligned}\rho D_0 + \rho D_1\tilde{Q} + \rho D_2\tilde{Q}^2 = &\frac{1}{2\alpha}(1-D_1^2) + \overline{wh}(1-\frac{1}{\xi}n^*)D_1 + \frac{p^c}{\alpha}D_1 - \frac{(p^cD_1)^2}{\alpha}\\ &+\left[rD_1 + \frac{p^cD_1D_2}{\alpha} + 2\left(\frac{p^c}{\alpha}D_2 - \frac{(p^c)^2}{\alpha}D_1D_2\right) + 2\overline{wh}(1-\frac{1}{\xi}n^*)D_2\right]\tilde{Q}\\ &+\left[2rD_2 + \frac{6(p^cD_2)^2}{\alpha}\right]\tilde{Q}^2.\end{aligned} \tag{101}$$

Thereby we arrive at three equations for the three unknowns D_0, D_1 and D_2. Equating the coefficients of $\tilde{Q}^2$ on left and right-hand sides of (101) yields D_2 as described in (70).[8] Similarly Equating the coefficients of $\tilde{Q}$ on both sides the equality in $\tilde{Q}$ yields D_1 as expressed in (71). One can thereafter derive the expression for D_0 as well, which is given by

$$D_0 = \frac{1}{\rho}\left(\frac{1}{2\alpha}(1-D_1^2) + \overline{wh}(1-\frac{1}{\xi}n^*)D_1 + \frac{p^c}{\alpha}D_1 - \frac{(p^cD_1)^2}{\alpha}\right). \tag{102}$$

□

Not all agents will invest in the whole index spectrum of human capital $[I_t^*, N_t^*]$. In fact, as illustrated later on, only measure zero of agents in the continuum will do so. Hence, we derive corresponding decision rules for agents who only invest in a subset of indexes.

Type 2 - Specializers: Specializers have manageable human-capital-augmented wage rate on a subset of the index domain, i.e. $w_{jt}h_{ijt} = w_{j't}h_{ij't}$, for all $j, j' \in \mathcal{D} \subseteq [I_t^*, N_t^*]$ while $w_{jt}h_{ijt} > w_{kt}h_{ikt}$, for all $j \in \mathcal{D}$ and for all $k \in \mathcal{D}^c$.[9]

Proposition A.3. *Assume an agent has manageable human-capital-augmented wage rate on a subset of the index domain $\mathcal{D}$ with measure $\nu(\mathcal{D})$, where $\nu : \mathcal{F}([I_t^*, N_t^*]) \mapsto \mathbb{R}_+$ is a measure function such that $\nu(\emptyset) = 0$ and $\nu([I_t^*, N_t^*]) = n_t^*$ and $\mathcal{F}([I_t^*, N_t^*])$ is the smallest possible sigma-algebra defined on $[I_t^*, N_t^*]$. Then under assumptions 3 and 4'', the agent facing the problem in (34) subject to (36), (39), (16), (17) and (41), is a specializer on $\mathcal{D}$ and acts in accordance with the decision rules in (106) and (107) but only for $j \in \mathcal{D}$. Moreover, every expression including integration over $[I_t^*, N_t^*]$ is replaced by integration over $\mathcal{D}$ and n_t^* is replaced by $\nu(\mathcal{D})$. Finally, in all expressions, the average human-capital-augmented wage rate $\overline{wh}_{it}$ is replaced by*

$$\overline{wh}_{it}^{\mathcal{D}} \equiv \frac{1}{\nu(\mathcal{D})}\int_{\mathcal{D}} w_{j,t}h_{ijt}dj. \tag{103}$$

[8] The equality also yields $D_2 = 0$, which results in $D_0 = D_1 = 0$ and hence $V \equiv 0$ the trivial zero solution to the PDE.

[9] Observe that $\mathcal{D}^c$ is the complement set of $\mathcal{D}$, i.e. $\mathcal{D}^c = [I_t^*, N_t^*] \setminus \mathcal{D}$.

Proof. The proof is the same as for Propositions A.2 and A.4 but with the adjustments mentioned in the text of Proposition A.3 above. $\square$

Scenario 2 - Only Asset Volatility: We now derive the decision rules for the case with following structure on asset volatility.

Assumption 4″.

$$\textit{No jump risk: } \lambda = 0, \textit{ but asset risk given by } \sigma_i(x_{it}, \boldsymbol{h}_{it}) = \sigma\sqrt{2x\left(\widetilde{Q} + \frac{D_{1t}}{2D_{2t}}\right)},$$

where $\widetilde{Q}$ is the same as in (69), but D_{ijt}, D_{1t} and D_{2t} are instead given by,

$$D_{ijt} = \sqrt{\frac{2e^{\gamma_g j}\overline{wh}_{it}w_{ijt}}{4\xi(r_t+\sigma^2) - e^{-\gamma_g j}}}, \qquad D_{2t} = \frac{\alpha(\rho - 2(r_t+\sigma^2))}{6\,(p_t^c)^2} \tag{104}$$

$$D_{1t} = \frac{(\rho - 2(r_t+\sigma^2))\left[1 + \frac{\alpha\overline{wh}_{it}(1-\frac{1}{\xi}n_t^*)}{p_t^c}\right]}{p_t^c\,(4\rho - 5(r_t+\sigma^2))}. \tag{105}$$

Type 1 - Diversifiers:

Proposition A.4. *Under assumptions 3 and 4″, the agent with manageable human-capital-augmented wage rate facing the problem in (34) subject to (36), (39), (16), (17) and (41), is a diversifier and acts in accordance with the following decision rules:*

$$a_{ijt}^* = \frac{1}{\xi}\left(1 - \sqrt{(4e^{\gamma_g j}\xi(r_t+\sigma^2)-1)\frac{w_{jt}h_{ijt}}{\overline{wh}_{it}}}\right), \qquad \ell_{ijt}^* = \frac{1}{n_t^*} - a_{ijt}^* \tag{106}$$

$$c_{it} = \frac{1}{\alpha}\left(1 - D_{1it} - D_{2t}\widetilde{Q}\right), \tag{107}$$

where $\widetilde{Q}$ is the same as in (69), but D_{ijt}, D_{1t} and D_{2t} are instead given by (104) and (105).

Proof of Proposition A.4. As done previously, for ease of notation we suppress the agent and time indexes i and t. The calculations up to the HJB-equation (90) are the same as in the proof of Proposition A.2, which then by Assumption 4″ is reduced to the following PDE,

$$\begin{aligned}\rho V = {} & \frac{1}{2\alpha}\left(1 - \left(p^c\frac{\partial V}{\partial x}\right)^2\right) + \frac{\partial V}{\partial x}\left[rx + \overline{wh}\left(1 - \frac{1}{\xi}\left(n^* - \frac{\partial V}{\partial x}\int_{I^*}^{N}\frac{e^{\gamma_g j}w_j}{\frac{\partial V}{\partial h_j}}dj\right)\right) - \frac{p^c}{\alpha}\left(1 - p^c\frac{\partial V}{\partial x}\right)\right] \\ & + \frac{1}{2\xi}\int_{I^*}^{N}\left(e^{-\gamma_g j}h_j\frac{\partial V}{\partial h_j} - \overline{wh}w_j e^{\gamma_g j}\frac{\left(\frac{\partial V}{\partial x}\right)^2}{\frac{\partial V}{\partial h_j}}\right)dj + \sigma^2 x\left(2D_2\tilde{Q} + D_1\right).\end{aligned} \tag{108}$$

To solve this problem we employ the structure of *ansatz* as in (92) for the optimal current-value function. We derive then (94) to (99) in the same manner. One difference is that instead of (100) we have the following result:

$$\frac{\overline{wh}e^{\gamma_g j}w_j}{2\xi D_j} + \frac{e^{-\gamma j}D_j}{4\xi} = (r+\sigma^2)D_j. \tag{109}$$

Inserting (92) and (94)-(99) into the PDE (108) and employing (109) yields:

$$\begin{aligned}\rho D_0 + \rho D_1\tilde{Q} + \rho D_2\tilde{Q}^2 = \frac{1}{2\alpha}(1-D_1^2) + \overline{wh}(1-\frac{1}{\xi}n^*)D_1 + \frac{p^c}{\alpha}D_1 - \frac{(p^cD_1)^2}{\alpha} \\ + \left[(r+\sigma^2)D_1 + \frac{p^cD_1D_2}{\alpha} + 2\left(\frac{p^c}{\alpha}D_2 - \frac{(p^c)^2}{\alpha}D_1D_2\right) + 2\overline{wh}(1-\frac{1}{\xi}n^*)D_2\right]\tilde{Q} \\ + \left[2(r+\sigma^2)D_2 + \frac{6(p^cD_2)^2}{\alpha}\right]\tilde{Q}^2. \end{aligned} \tag{110}$$

Thereby we arrive at three equations for D_0, D_1 and D_2. Equating the coefficients of $\tilde{Q}^2$ and $\tilde{Q}$ on left and right-hand sides of (101) yields the tautology $0 = 0$, validating the *ansatz*.[10] One can thereafter derive the expression for D_0 as well, which is given by (102). □

Type 2 - Specializers:

Proposition A.5. *Assume an agent has manageable human-capital-augmented wage rate on a subset of the index domain $\mathcal{D}$ with measure $\nu(\mathcal{D})$, where $\nu : \mathcal{F}([I_t^*, N_t^*]) \mapsto \mathbb{R}_+$ is a measure function such that $\nu(\emptyset) = 0$ and $\nu([I_t^*, N_t^*]) = n_t^*$ and $\mathcal{F}([I_t^*, N_t^*])$ is the smallest possible sigma-algebra defined on $[I_t^*, N_t^*]$. Then under assumptions 3 and 4'', the agent facing the problem in (34) subject to (36), (39), (16), (17) and (41), is a specializer on $\mathcal{D}$ and acts in accordance with the decision rules in (106) and (107) but only for $j \in \mathcal{D}$. Moreover, every expression including integration over $[I_t^*, N_t^*]$ is replaced by integration over $\mathcal{D}$ and n_t^* is replaced by $\nu(\mathcal{D})$. Finally, in all expressions, the average human-capital-augmented wage rate $\overline{wh}_{it}$ is replaced by $\overline{wh}_{it}^{\mathcal{D}}$.*

Proof. The proof is the same as for Propositions A.4 but with the adjustments mentioned in the text of Proposition A.5 above. □

Scenario 3 - With Both Asset and Obsolescence Risks: We derive the following two lemmas. Lemma A.2 provides conditions for diversifiers and specializers, while Lemma A.3 derives a lower bound for the agents optimal current-value loss under obsolescence. Hence, by employing this lower bound in the derivation of agents' decision rules we are assuming that the behavior is slightly more cautious than optimal and therefore formally are finding a so called *supersolution* to the optimization problem.

Lemma A.2. *Under positive jump risk $\lambda > 0$, the agent is a specializer if and only if she has manageable human-capital-augmented wage rate such that*

$$w_{jt}h_{ijt}F_{I_t+\Delta_{I_t}}(j) = w_{j't}h_{ij't}F_{I_t+\Delta_{I_t}}(j') \text{ for all } j, j' \in [I_t^*, N_t^*]. \tag{111}$$

Moreover, the agent is a diversifier if and only if she has manageable human-capital-augmented wage rate on a subset of the index domain $\mathcal{D} \subseteq [I_t^*, N_t^*]$ such that

$$\begin{aligned} w_{jt}h_{ijt}F_{I_t+\Delta_{I_t}}(j) = w_{j't}h_{ij't}F_{I_t+\Delta_{I_t}}(j') \text{ for all } j, j' \in \mathcal{D}, \\ \text{while } w_{jt}h_{ijt}F_{I_t+\Delta_{I_t}}(j) > w_{kt}h_{ikt}F_{I_t+\Delta_{I_t}}(k), \text{ for all } j \in \mathcal{D} \text{ and for all } k \in \mathcal{D}^c. \end{aligned} \tag{112}$$

[10] As before, $V \equiv 0$ is the trivial zero solution to the PDE.

Proof of Lemma A.2. First consider the condition for agents having manageable human-capital-augmented wage rate without any jump risk $\lambda = 0$

$$w_{jt}h_{ijt} = w_{j't}h_{ij't}$$

we rewrite it as

$$w_{jt}h_{ijt^-} \cdot e^{g_j(a^*_{ijt})dt} = w_{j't}h_{ij't^-} \cdot e^{g_{j'}(a^*_{ij't})dt} \tag{113}$$

where

$$t^- = \lim_{\epsilon \downarrow 0, \epsilon > 0} t - \epsilon. \tag{114}$$

With jump risk the condition for agents having manageable human-capital-augmented wage rate

$$\begin{aligned} &w_{jt}h_{ijt^-} \cdot e^{g_j(a^*_{ijt})dt} \cdot 1_{\{h_{ij}(t^- + dt) \neq 0 | h_{ijt^-} \neq 0\}} \\ &\qquad = w_{j't}h_{ij't^-} \cdot e^{g_{j'}(a^*_{ij't})dt} \cdot 1_{\{h_{ij'}(t^- + dt) \neq 0 | h_{ij't^-} \neq 0\}} \end{aligned} \tag{115}$$

Observe that

$$1_{\{h_{ij}(t^- + dt) \neq 0 | h_{ijt^-} \neq 0\}} = 1_{\{T_{I_t} > t^- + dt | T_{I_t} > t^-\}} \cdot 1_{\{I_t \Delta_{I_t} < j\}}. \tag{116}$$

where T_{I_t} is the stochastic variable for time of new automation technology arriving. Moreover we have that,

$$\mathbb{E}\left\{1_{\{T_{I_t} > t^- + dt | T_{I_t} > t^-\}}\right\} = \frac{\mathcal{P}\left(T_{I_t} > t^- + dt\right)}{\mathcal{P}\left(T_{I_t} > t^-\right)} = \frac{e^{\lambda(t^- + dt)}}{e^{\lambda t^-}} = e^{\lambda dt} \tag{117}$$

where we have used that the distribution of a arrival time of Poisson processes $Poi(\lambda(t))$ are given by Exponential distributions $Exp(\lambda(t))$. Furthermore we have that,

$$\mathbb{E}\left\{1_{\{I_t + \Delta_{I_t} < j\}}\right\} = F_{I_t + \Delta_{I_t}}(j) \tag{118}$$

assuming continuity of density $f_{\Delta_{I_t}}(j)$ over the support $j \in [N_t, I^*_t]$. Thus by (116)-(118), taking expectation on both sides of (115) and comparing the result to (113) , the condition for agents having manageable human-capital-augmented wage rate becomes (111). Finally, (112) follows from correspondingly similar calculations as above. □

Lemma A.3. *Assume the optimal value-function has the same structure as the* ansatz *in (92). Then, under positive jump risk $\lambda > 0$, the following holds*

$$\lambda \left[\int_0^{1-n^*_t} (V(., N + s) - V) f_{\Delta_{N^*_t}}(s) ds\right] \geq -\lambda \mathcal{P}_{\Delta_{N^*_t}} x_{it}(D_{1t} + 2D_{2t}\widetilde{Q}), \tag{119}$$

where

$$\mathcal{P}_{\Delta_{N_t^*}} \equiv \mathcal{P}_{[N_t^*-1, N_t^*-1+\mathbb{E}\,\Delta_{N_t^*}],t}. \tag{120}$$

Proof. When new tasks arrive, the portion of assets given by (120) is going to become obsolete. Hence, by (92) and (69), then we have that

$$\int_0^{1-n_t^*} (V(.,N+s)-V) f_{\Delta_{N_t^*}}(s)ds = -\mathcal{P}_{\Delta_{N_t^*}} x_{it} \left(D_{1t} + 2D_{2t} \left(\mathcal{P}_{\Delta_{N_t^*}} x_{it} + \int_{I_t^*}^{N_t^*} D_{ijt}\sqrt{2h_{ijt}}dj \right)\right). \tag{121}$$

Since $0 \le \mathcal{P}_{\Delta_{N_t^*}} \le 1$, (119) follows promptly. □

Type 1 - Diversifiers:

Proposition A.6. *Under assumptions 3 and 4, the agent with manageable human-capital-augmented wage rate facing the problem in (34) subject to (36), (39), (16), (17) and (41), is a diversifier and acts in accordance with the following decision rules:*

$$a_{ijt}^* = \frac{1}{\xi}\left(1 - \sqrt{(4e^{\gamma_g j}\xi(r_t + \sigma^2 - \lambda \mathcal{P}_{\Delta_{N_t^*}}) - 1)\frac{w_{jt}h_{ijt}F_{\Delta_{I_t}}(j)}{\overline{wh}_{it}}}\right), \quad \ell_{ijt}^* = \frac{1}{n_t^*} - a_{ijt}^* \tag{122}$$

$$c_{it}^* = \frac{1}{\alpha}\left(1 - D_{1it} - D_{2t}\widetilde{Q}\right), \tag{123}$$

where $\widetilde{Q}$ is the same as in (69), but D_{ijt}, D_{1t} and D_{2t} are instead given by (45), (46) and (47).

Proof of Proposition A.6. As done previously, for ease of notation we suppress the agent and time indexes i and t. The calculations up to the HJB-equation (90) are the same as in the proof of Proposition A.2, which then by Assumption 4 and Lemma A.2 is reduced to the following PIDE,

$$\begin{aligned}\rho V = {} & \frac{1}{2\alpha}\left(1 - \left(p^c \frac{\partial V}{\partial x}\right)^2\right) + \frac{\partial V}{\partial x}\left[rx + \overline{wh}\left(1 - \frac{1}{\xi}\left(n^* - \frac{\partial V}{\partial x}\int_{I^*}^{N} \frac{e^{\gamma_g j} w_j F_{\Delta_{I_t}}(j)}{\frac{\partial V}{\partial h_j}} dj\right)\right) - \frac{p^c}{\alpha}\left(1 - p^c \frac{\partial V}{\partial x}\right)\right] \\ & + \frac{1}{2\xi}\int_{I^*}^{N}\left(e^{-\gamma_g j} h_j \frac{\partial V}{\partial h_j} - \overline{wh} w_j F_{\Delta_{I_t}}(j) e^{\gamma_g j} \frac{\left(\frac{\partial V}{\partial x}\right)^2}{\frac{\partial V}{\partial h_j}}\right) dj + \sigma^2 x \left(2D_2\tilde{Q} + D_1\right) \\ & + \lambda\left[\int_0^{1-n_t^*} (V(.,N+s) - V) f_{\Delta_{N_t}}(s) ds\right].\end{aligned} \tag{124}$$

To solve this problem we employ the structure of *ansatz* as in (92) for the optimal current-value function but D_{ijt}, D_{1t} and D_{2t} are instead given by (45), (46) and (47). Employing (93), we get

$$\frac{1}{2\alpha}\left(1-\left(p^c\frac{\partial V}{\partial x}\right)^2\right)=\frac{1}{2\alpha}(1-D_1^2)+\frac{p^cD_1D_2}{\alpha}\tilde{Q}+\frac{2(p^cD_2)^2}{\alpha}\tilde{Q}^2 \tag{125}$$

$$rx\frac{\partial V}{\partial x}=rD_1x+2rD_2x\tilde{Q} \tag{126}$$

$$\begin{aligned}\frac{\partial V}{\partial x}\overline{wh}\left(1-\frac{1}{\xi}\left(n^*-\frac{\partial V}{\partial x}\int_{I^*}^{N}\frac{e^{\gamma_g j}w_jF_{\Delta_I}(j)}{\frac{\partial V}{\partial h_j}}dj\right)\right)&=\overline{wh}(1-\frac{1}{\xi}n^*)D_1+2\overline{wh}(1-\frac{1}{\xi}n^*)D_2\tilde{Q}\\&+\frac{\overline{wh}}{\xi}D_1\int_{I^*}^{N}\frac{e^{\gamma_g j}w_j}{D_j}\sqrt{2h_j}dj+2\frac{\overline{wh}}{\xi}D_2\int_{I^*}^{N}\frac{e^{\gamma_g j}w_jF_{\Delta_I}(j)}{D_j}\sqrt{2h_j}dj\tilde{Q}\end{aligned} \tag{127}$$

$$\frac{\partial V}{\partial x}\frac{p^c}{\alpha}\left(1-p^c\frac{\partial V}{\partial x}\right)=\frac{p^c}{\alpha}D_1-\frac{(p^cD_1)^2}{\alpha}+2\left(\frac{p^c}{\alpha}D_2-\frac{(p^c)^2}{\alpha}D_1D_2\right)\tilde{Q}-\frac{4\,(p^cD_2)^2}{\alpha}\tilde{Q}^2 \tag{128}$$

$$\frac{1}{2\xi}\int_{I^*}^{N}e^{-\gamma_g j}h_j\frac{\partial V}{\partial h_j}dj=\frac{D_1}{2\xi}\int_{I^*}^{N}\frac{e^{-\gamma_g j}D_j}{2}\sqrt{2h_j}dj+\frac{D_2}{\xi}\int_{I^*}^{N}\frac{e^{-\gamma_g j}D_j}{2}\sqrt{2h_j}dj\tilde{Q} \tag{129}$$

$$\frac{1}{2\xi}\int_{I^*}^{N}\overline{wh}w_jF_{\Delta_{I_t}}(j)e^{\gamma_g j}\frac{\left(\frac{\partial V}{\partial x}\right)^2}{\frac{\partial V}{\partial h_j}}dj=\frac{D_1}{2\xi}\int_{I^*}^{N}\frac{\overline{wh}w_jF_{\Delta_{I_t}}(j)e^{\gamma_g j}}{D_j}\sqrt{2h_j}dj \tag{130}$$

$$+\frac{D_2}{\xi}\int_{I^*}^{N}\frac{\overline{wh}w_jF_{\Delta_{I_t}}(j)e^{\gamma_g j}}{D_j}\sqrt{2h_j}dj\tilde{Q}. \tag{131}$$

Observe also that

$$\frac{\overline{wh}e^{\gamma_g j}w_jF_{\Delta_{I_t}}(j)}{2\xi D_j}+\frac{e^{-\gamma j}D_j}{4\xi}=(r+\sigma^2-\lambda\mathcal{P}_{\Delta_{N_t}})D_j. \tag{132}$$

Inserting (92) and (125)-(131) into the PDE (124) and employing (132) and Lemma (A.3) yields:

$$\begin{aligned}\rho D_0+\rho D_1\tilde{Q}+\rho D_2\tilde{Q}^2&=\frac{1}{2\alpha}(1-D_1^2)+\overline{wh}(1-\frac{1}{\xi}n^*)D_1+\frac{p^c}{\alpha}D_1-\frac{(p^cD_1)^2}{\alpha}\\&+\left[(r+\sigma^2-\lambda\mathcal{P}_{\Delta_{N_t}})D_1+\frac{p^cD_1D_2}{\alpha}+2\left(\frac{p^c}{\alpha}D_2-\frac{(p^c)^2}{\alpha}D_1D_2\right)+2\overline{wh}(1-\frac{1}{\xi}n^*)D_2\right]\tilde{Q}\\&+\left[2(r+\sigma^2-\lambda\mathcal{P}_{\Delta_{N_t}})D_2+\frac{6(p^cD_2)^2}{\alpha}\right]\tilde{Q}^2.\end{aligned} \tag{133}$$

Observe that we have assumed that agents are acting slightly more conservative by inserting the lower bound given in Lemma A.3. Thereby we arrive at three equations for D_0, D_1 and D_2. Equating the coefficients of $\tilde{Q}^2$ and $\tilde{Q}$ on left and right-hand sides of (133) yields the tautology $0=0$, validating the *ansatz*.[11] One can thereafter derive the expression for D_0 as well, which is given by (102). □

Type 2 - Specializers:

Proposition A.7. *Assume an agent has manageable human-capital-augmented wage rate on a subset of the index domain $\mathcal{D}$ with measure $\nu(\mathcal{D})$, where $\nu:\mathcal{F}([I_t^*,N_t^*])\mapsto\mathbb{R}_+$ is a measure function such that $\nu(\emptyset)=0$ and $\nu([I_t^*,N_t^*])=n_t^*$ and $\mathcal{F}([I_t^*,N_t^*])$ is the smallest possible σ-algebra defined on $[I_t^*,N_t^*]$. Then under assumptions 3 and 4, the agent facing the problem in (34) subject to (36), (39), (16), (17) and (41), is a specializer on $\mathcal{D}$ and acts in accordance with*

[11] As before, $V\equiv 0$ is the trivial zero solution to the PIDE.

the decision rules in (122) and (123) but only for $j \in \mathcal{D}$. Moreover, every expression including integration over $[I_t^*, N_t^*]$ is replaced by integration over $\mathcal{D}$ and n_t^* is replaced by $\nu(\mathcal{D})$. Finally, in all expressions, the average human-capital-augmented wage rate $\overline{wh}_{it}$ is replaced by (42).

Proof. The proof is the same as for Propositions A.6 but with the adjustments mentioned in the text of Proposition A.7 above. □

Proof of Proposition 3.6. The proof follows from Propositions A.2 to A.7. □

Proof of Proposition 3.8. By (49) and (25) plus augmenting the learning function with the state-enhancement productivity investment $A(TX_t)$ we get the following

$$a_{ijt}^* \sim \frac{1}{\xi}\left(1 - \sqrt{e^{(\gamma_g + \left(\frac{\varepsilon-1}{\varepsilon}\right)\gamma_L + \delta_I - \gamma_{h_i})j}(4\xi(r_t + \sigma^2 - \lambda \mathcal{P}_{\Delta_{N_t^*}}) - A(TX_t)e^{-\gamma_g j})P_t \left(\frac{Y_t}{h_{jt}\ell_{jt}}\right)^{\frac{1}{\epsilon}} \frac{h_0}{\overline{wh}_{it}}}\right). \quad (134)$$

which yields the threshold $\widetilde{\gamma}_{h_i}$ in (51) under the conditions stated in the proposition. Observe that arrival rate of new tasks N_t is constant at rate λ, X_t grows at a constant rate on the BGP and $A' > 0, A'' < 0$. Then, augmented learning efficiency at the technological frontier $A(TX_t)e^{-\gamma_g N_t}$ will become small enough when $N_t > \underline{N}$ for some $\underline{N} > 0$, such that $A(TX_t)e^{-\gamma_g N_t} < 4\xi\epsilon$. Then, workers with higher endowment parameter $\gamma_{h_i} \geq \widetilde{\gamma}_{h_i}$ will have decreasing investment a_{ijt} along the task index j on their corresponding manageable human-capital-augmented wage rate on a subset of the index domain $\mathcal{D}_i = [I_t^*, \iota_i], \iota_i \in [I_t^*, N_t^*]$. Similarly, workers with lower endowment parameter $\gamma_{h_i} \leq \widetilde{\gamma}_{h_i}$ will have increasing investment a_{ijt} along the task index j on their corresponding manageable human-capital-augmented wage rate on the subset $\mathcal{D}_i = [\iota_i, N_t^*]$. □

A.2.1 Aggregation

In mean-field systems such as these the aggregate states are given by

$$Z_t = \int z_{it} di = \int_{-\infty}^{\infty} z_t \phi(z_t) dz_t$$

where z_{it} is any state variable and $\phi(z_t)$ is the transitional density given by a corresponding Kolmogorov-forward equation.[12] In this model, the perceived density is given by,

$$\begin{aligned}\frac{\partial}{\partial t}\phi(x, \boldsymbol{h}, t) &= (2p_t^c D_{2,t} - r_t)\,\phi(x, \boldsymbol{h}, t) - \frac{\partial}{\partial x}\phi(x, \boldsymbol{h}, t)\left(r_t x_t + \int_{I_t^*}^{N_t^*} w_{jt} h_{jt} \ell_{jt}^*(x, \boldsymbol{h}) dj\right) \\ &- \frac{p_t^c}{\alpha}\left[1 - D_{1,t} - 2D_{2,t}\widetilde{Q}\right] - \int_{I_t^*}^{N_t^*} g_j(a_j^*(x, \boldsymbol{h}))\phi(x, \boldsymbol{h}, t) dj - \int_{I_t^*}^{N_t^*} g_j(a_j^*(x, \boldsymbol{h}))\frac{\partial}{\partial h_j}\phi(x, \boldsymbol{h}, t) h_{jt} dj \\ &\frac{1}{2}\left[2\sigma^2\phi(x, \boldsymbol{h}, t) + 2\sigma^2\left(\widetilde{Q} + x_t + \frac{D_{1,t}}{2D_{2,t}}\right)\frac{\partial}{\partial x}\phi(x, \boldsymbol{h}, t) + \frac{\sigma^2}{2}\left(x_t(\widetilde{Q} + \frac{D_{1,t}}{2D_{2,t}})\right)\frac{\partial^2}{\partial x^2}\phi(x, \boldsymbol{h}, t)\right]\end{aligned} \quad (135)$$

In equilibrium $\frac{\partial}{\partial t}\phi(x, \boldsymbol{h}) = 0$. This density is of course conditional to the initial distribution on human capital endowment. The actual joint density $\psi(x, \boldsymbol{h}, t)$ is given as the product of the

[12] Also known as the Fokker-Planck equation in the physics literature.

density $\phi(x, \boldsymbol{h}, t)$ and $s(\gamma_h; \varrho)$.

$$\psi(x, \boldsymbol{h}, t, \gamma_h; \varrho) = \phi(x, \boldsymbol{h}, t) s(\gamma_h; \varrho) \tag{136}$$

where $s(\gamma_h; \varrho)$ is the density of an inverse exponential distribution with parameter $\varrho > 0$.

B Figures

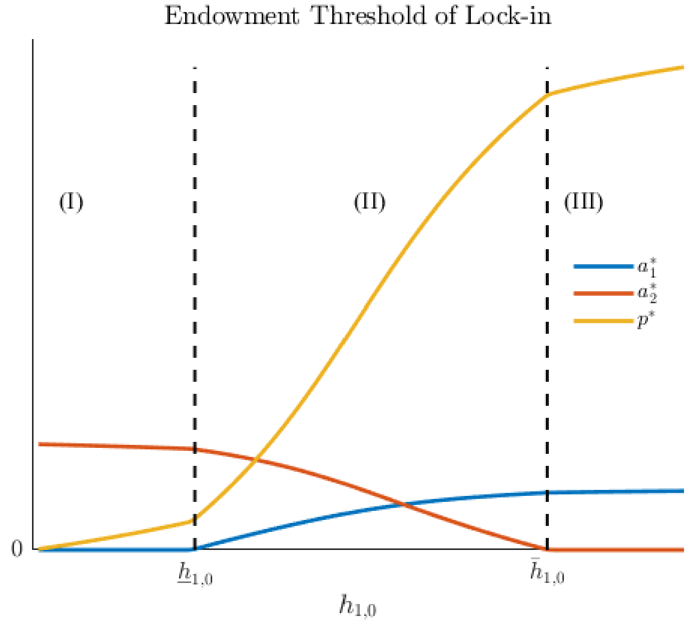

Figure 1: Example with Diversifiers.

The figure shows optimal attention-to-learning skills 1 and 2 ($a_j^*, j \in \{1, 2\}$) and skill portfolio p^* as function of human capital endowment in skill 1 $h_{1,0}$ keeping stock in skill 2 constant. Human capital endowment thresholds of human capital in skill 1 ($\underline{h}_{1,0}$ and $\bar{h}_{1,0}$) for being locked into different types of workers (I) specializers in task 2, (II) diversifiers and (III) specializers in task 1 are marked with a dashed line.

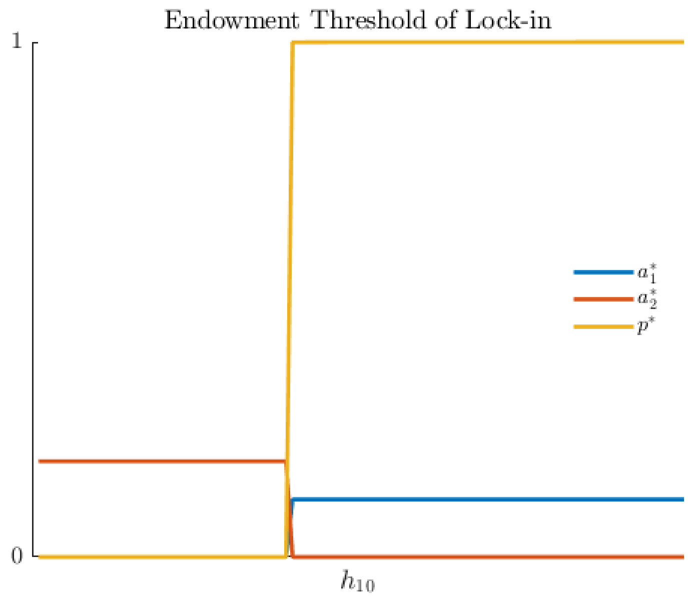

Figure 2: Example with no Diversifiers.
The figure shows optimal attention-to-learning skills 1 and 2 ($a_j^*, j \in \{1, 2\}$) and skill portfolio p^* as function of human capital endowment in skill 1 $h_{1,0}$ keeping stock in skill 2 constant. This example illustrates the result of Proposition 3.3, which states that when there is no risk for obsolescence, there are no diversifiers and all workers specialize in either skill depending on their initial endowments.

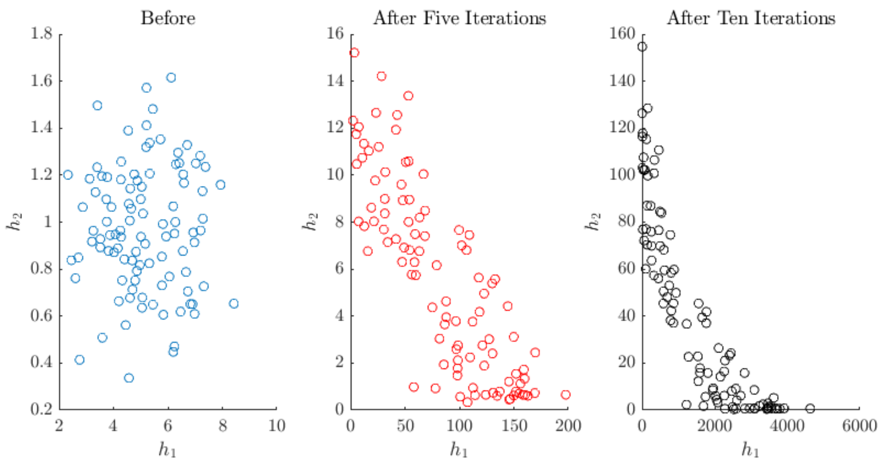

Figure 3: Intergenerational Transmission of Capital in an Example with Diversifiers.
Each circle shows a worker dynasti with different stocks of human capital in skills 1 h_1 and 2 h_2. The leftmost subfigure shows the initial endowment structure in the population. The next two subfigures show the resulting changes after five and ten iterations of the optimization problem. Each iteration corresponds to one generation. The examples are calculated with $U(C) = 2C^{0.5}$ $g_1(a) = 2 - \exp(-20a)$, $g_2(a) = 2 - \exp(-5a)$, $\xi_1 = 0.6$ and $\xi_2 = 0.5$.

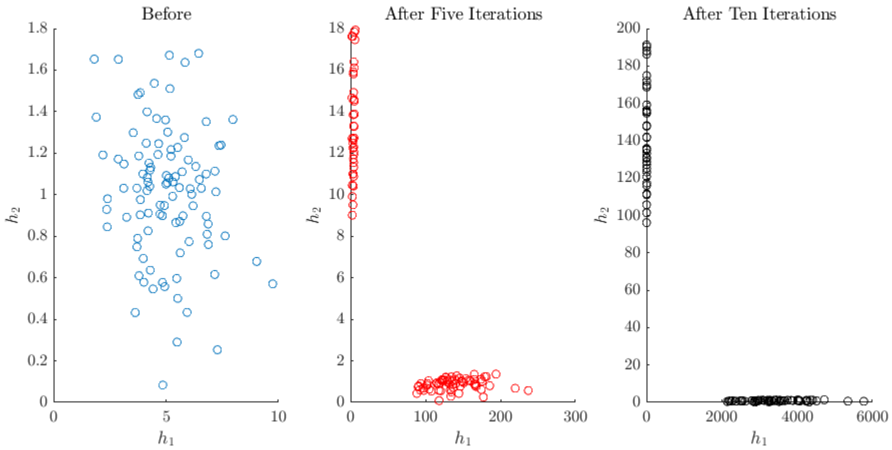

Figure 4: Intergenerational Transmission of Capital in an Example with no Diversifiers.
Each circle shows a worker dynasti with different stocks of human capital in skills 1 h_1 and 2 h_2. The leftmost subfigure shows the initial endowment structure in the population. The next two subfigures show the resulting changes after five and ten iterations of the optimization problem. Each iteration corresponds to one generation. The examples are calculated with $U(C) = 2C^{0.5}$ $g_1(a) = 2 - \exp(-20a)$ and $g_2(a) = 2 - \exp(-5a)$ as in Figure 3, but with $\xi_1 = 0.01$ and $\xi_2 = 10^{-4}$, leading to no diversification in human capital profiles.

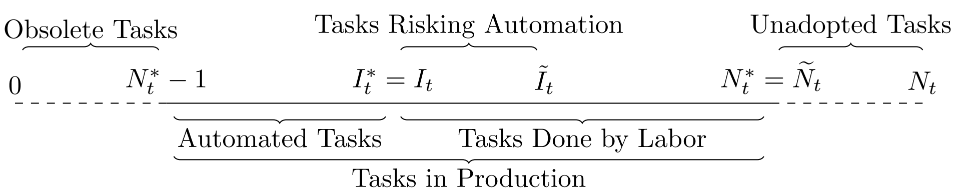

Figure 5: Task Spectrum in the Full Model.
Tasks are ordered from routine to non-routine. Tasks in production are in the set $[N_t^* - 1, N_t^*]$. Tasks below $N_t^* - 1$ are obsolete, and tasks beyond $\widetilde{N}_t$ are unadopted. I_t and N_t are the available automation technolgy and new tasks respectively. Here $N_t^* = \widetilde{N}_t$ and $I_t^* = I_t$. Hence tasks between available automation technolgy I_t and $\tilde{I}_t$ are risking automation. Tasks in $[N_t^* - 1, I_t^*]$ are done by machines and those in $[I_t^*, N_t^*]$ by labor.

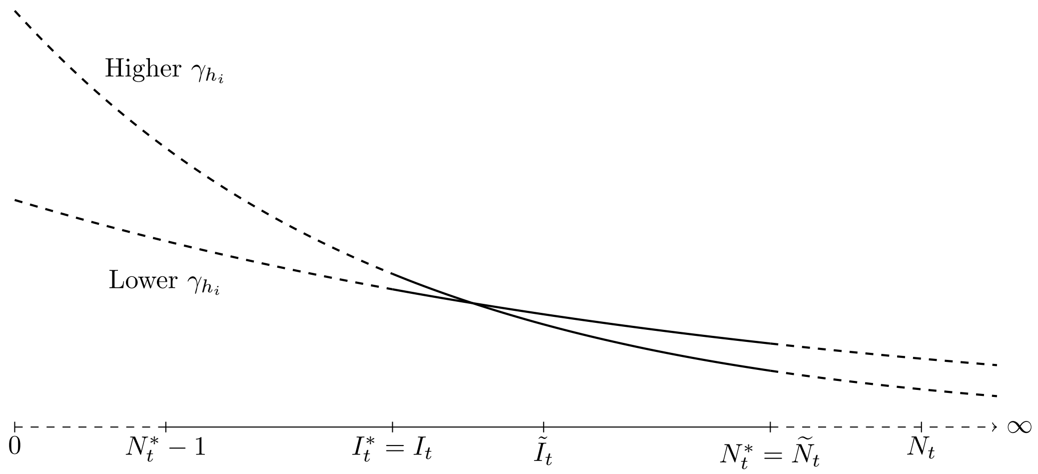

Figure 6: Endowment profiles in stock-density $f_{h_{i0}}(j)$ over the task spectrum j.
Endowment profiles are exponential $f_{h_{i0}}(j) \sim e^{-\gamma_{h_i} j}$ and depicted for two workers – one with higher and one with lower γ_{h_i}. The full line is the $f_{h_{i0}}(j)$ profiles for the tasks done labor $j \in [I_t^*, N_t^*]$, and the dashed portions depicts their analytic continuations.

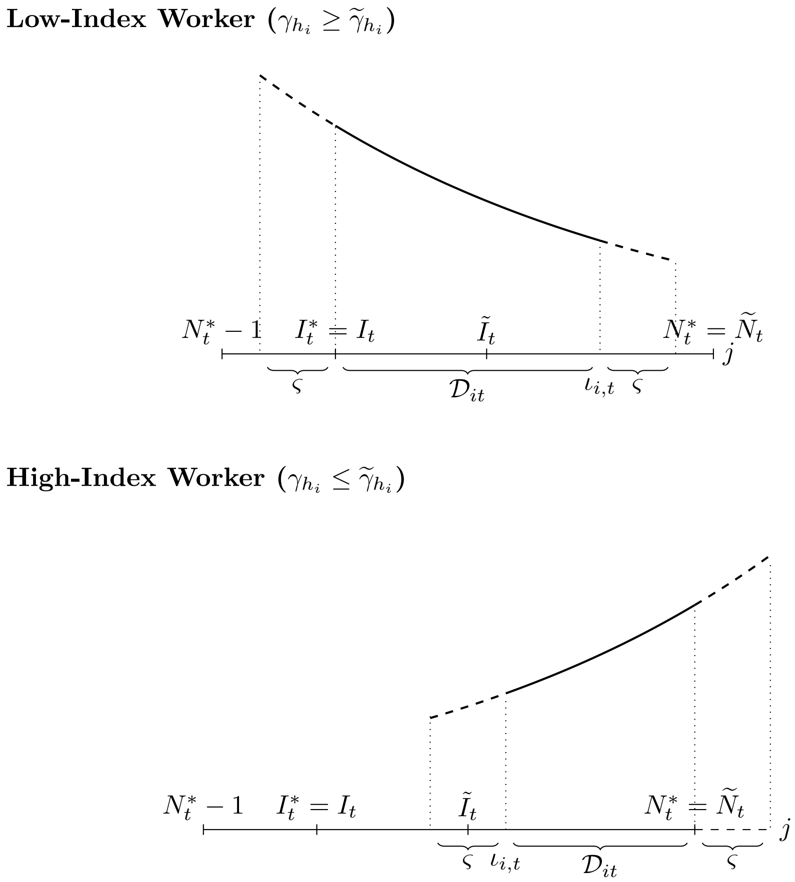

Figure 7: Human capital investment profiles a^*_{ijt} over the task spectrum j.

The figure shows the optimal attention-to-learning schemes for low and high-index workers. Low-index workers optimize over sets of the form $\mathcal{D}_{it} = [I_t^*, \iota_{i,t}]$, while for high-index workers the sets are of the form $\mathcal{D}_{it} = [\iota_{i,t}, N_t^*]$. The ful segments of the lines depicts a^*_{ijt} and the dashed segments the locally crossproductive regions with distance $\varsigma > 0$. Optimal attention to learning a^*_{ijt} for low-index workers is focused focused on more routine tasks, while for high-index worers learning non-routine tasks dominates.