Affording Superstardom: Explaining Skill Premia's Convexity in Education*

[Latest Version Here.]

Danial Ali Akbari†

This Version: October 2021

Abstract

I develop a theoretical framework that explains the increasing and convex pattern of skill premia through diminishing aversion towards the ambiguous possibility of skill obsolescence. High-income workers are shown to invest more in education as their concern with forgone income is progressively lower than their less credentialed counterparts. As a result, high-skill (low-skill) individuals invest in their stock of human capital beyond (below) what is optimal if the true obsolescence frequency was known to them. This learning glut (deficit) subsequently pays large dividends (losses) during unexpected episodes that exhibit increased ambiguity. A calibration of the model is able to match the skill premium curve in the US economy.

Keywords: Ambiguity aversion, skill premium, human capital accumulation, income inequality
JEL Codes: D31, E24, E70, J24

*I am deeply grateful for detailed comments on this paper from Markus Poschke and Tobias Broer. Moreover, I thank my supervisors Thomas Fischer, Emiliano Santoro and Joakim Westerlund for support and helpful insights during the process of writing this paper. I am also thankful for helpful comments from a large number of researchers. In no particular order, these include Zsófia Bárány, Søren Hove Ravn, Kaveh Majlesi, Alessandro Martinello, Erik Mohlin, Morten Olsen, Pascual Restrepo, Alexandros Rigos, Petr Sedláček, Raman Uppal, participants in the Lund University macro-metric and theoretical microeconomics seminars, participants in the Copenhagen Macro seminars, researchers at Department of Economics at Copenhagen Business School that attended a presentation of this paper 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 (danial.ali_akbari@nek.lu.se).

†School of Economics and Management, Department of Economics, Knut Wicksell Center for Financial Studies and Center for Economic Demography at Lund University, Lund, Sweden.

1 Introduction

It is broadly documented that hourly wages and skill premia – defined as relative wage rates – are increasing and convex functions of schooling years (cf. e.g. Lemieux, 2006, 2008; Acemoglu and Autor, 2011; Autor, 2014). Additionally, the degree of said convexity has increased during the last three decades in the United States (*ibid.*). At the same time, the modern labor market has become increasingly uncertain and turbulent (Gottschalk and Moffitt, 2009; Lalé, 2018). As a result workers' skills can depreciate in usefulness and relevance – a process referred to as *skill obsolescence* – due to significant and sudden shifts in economic fundamentals. Trade deals can pave the way for firms to offshore and outsource production or services far away (Bhagwati, Panagariya and Srinivasan, 2004; Ebenstein et al., 2014). New technologies such as automation (Graetz and Michaels, 2018; Acemoglu and Restrepo, 2020) and digitization (Goos, Manning and Salomons, 2014; Hershbein and Kahn, 2018) can be introduced displacing employees into precariousness. Significant environmental shifts such as climate change may damage the productive capacities of the economy displacing workers in certain sectors directly or indirectly following measures to contain the deterioration (Millner, Dietz and Heal, 2013; Kahn et al., 2019). More recently, permanent job losses have followed pandemics due to pathogens with ambiguous attributes (Barrero, Bloom and Davis, 2020). In this study, I develop a theoretical framework where skill premia's convexity in schooling years, and the increased convexity in the last three decades, are explained through the particular pattern that emerges when workers make decisions on educational attainment while facing the uncertainty of skill obsolescence.

Facing uncertainty may understandably discourage skill acquisition. Imagine for instance a medical student who is considering radiology as a specialization but is also aware that image recognition software is being rapidly developed to identify lesions in the lung or cracks in the bone. She, however, does not know when the software will be ready for use or to which extent it will impact the prospects of her skills. In other words, skills can become obsolete but the timing and extent of obsolescence are uncertain. Consequently, returns to education – i.e. the process of acquiring skills – are perceived to be ambiguous. It is well-documented, moreover, that people are averse to ambiguous outcomes (cf. e.g. Ellsberg, 1961; Ahn et al., 2014; Baillon and Placido, 2019). In this study, I explore the interaction between ambiguity in prospects of further learning and employees aversion towards this uncertainty. I develop a theoretical framework that explains increasing and convex pattern of skill premia through said ambiguity and workers' aversion to it. I calibrate the model to US data matching patterns of skill premia and inequality. Moreover, I show that increased skill premium convexity in education observed in recent three decades in the United States is compatible with transitional episodes exhibiting increased ambiguity. These episodes include periods following events such as trade deals increasing import competition, labor-displacing and routine-biased technological change plus the climate crisis and its surrounding public policy discussion.

In the setting of this study, workers divide their attention between education and labor. High-income workers – i.e. those with already high human capital – tend to invest more in learning as their concern with forgone income due to skill obsolescence is progressively lower than their less credentialed counterparts. These affluent workers are subsequently able to reap the benefits of their exceedingly high educational attainments during transitional episodes of increased ambiguity, when less credentialed individuals try to keep up. Indeed, the affluence of high-income earners – some of which are "superstars" – affords them to optimally invest in their stock of human capital beyond the level which they would under the absence of ambiguity. Put differently, they would have invested less if they knew the true frequency of obsolescence. Correspondingly, low-skill and low-income earners will have a learning deficit. It is important to note that neither the low-skill nor the high-skill are aware of the fact that they are accumulating deficits and gluts in human capital respectively. The educational overinvestment by the high-skill leads to a learning glut which pays large dividends during unexpected transitions that exhibit increased ambiguity, all the while the rest of the population attempts to catch up in educational attainment. It is useful to highlight the key empirical observational findings that this paper attempts to explain. These are skill premia being increasing and mainly convex functions of schooling years, increased degree of convexity during the last three decades plus the accompanying patterns of inequality.

Figure 1 displays the development of hourly wage rates and pattern of skill premia over time. Subfigure (a) shows hourly wage rates (USD) of heads of households through their main source of employment. We see that wage rates are increasing in education. Moreover, in 1992 wages exhibited very little convexity, so that postgraduates (with 18 years of schooling) have almost the same hourly wage increase relative to college graduates (16 years) as individuals with some college graduation (14 years) would garner by completing their studies. In later years, however, wages and skill premia become increasingly convex functions of schooling especially at the top. This is well-illustrated in Subfigure (b) which shows hourly wage rates relative to that of college graduates in the same year. Skill premia of postgraduates increase while that of those with some college education fall over time. At lower end of schooling years – for high-school graduates (12 years) and dropouts (10 years) – skill premia remain constant, or even fall slightly. The large decrease in premia among those with some college education is so severe, however, that wage rates and skill premia appear to exhibit slightly concave behavior at the lower end of educational attainment. This severe fall in premia among some-college educated during the last three decades – documented well in Subfigure (c) – is followed by that of high-school dropouts and high-school graduates. This zigzag pattern of decrease highlights the slight concave pattern of wage rates and skill premia at the lower educational spectrum during the recent decades. The only group exhibiting wage growth are postgraduates. It is important to note that since the skill premium curve depicts relative wages, the values that it presents are readily comparable to the wage levels produced by the model, why skill premia is used in this study.

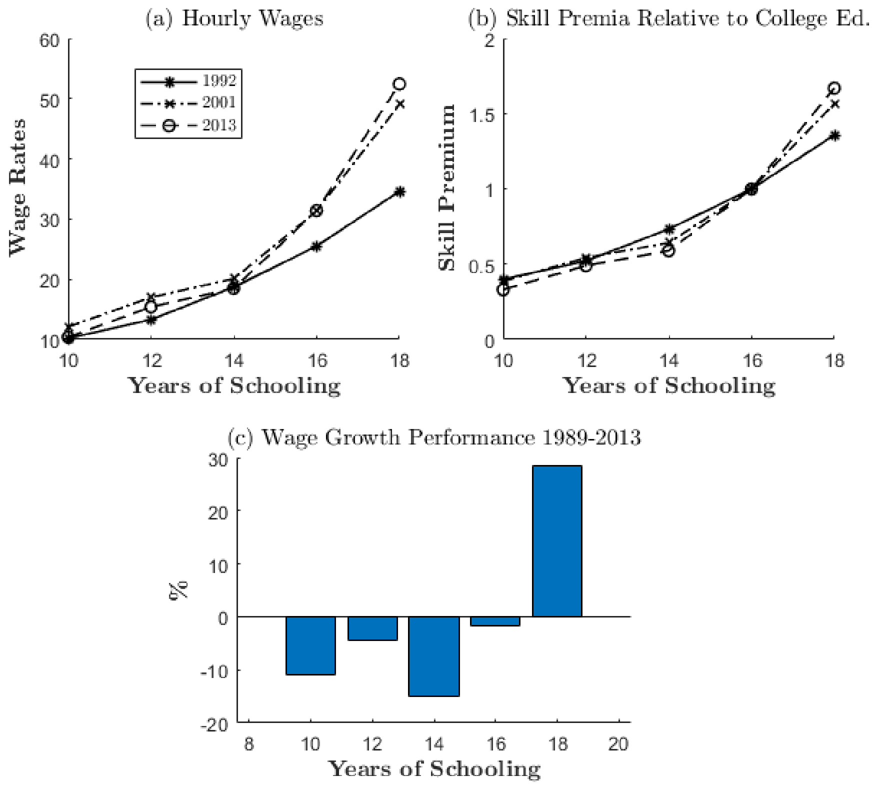

Figure 1: **Wages Rates and Skill Premia over Time.**

The education categories are highschool dropouts (10 years of schooling), highschool graduates (12 years), individuals with some college education (14 years), college graduates (16 years) and postgraduates (18 years). Subfigure (a) shows hourly wage rates (USD) of heads of households on their main source of employment for three years 1992, 2001 and 2013. Households with zero hours worked are excluded. Wages are computed as average labor income divided by average hours to reduce the eect of outliers on average wages. Source: Table 37 in Kuhn and Ríos-Rull (2016). Subfigure (b) shows hourly wage rates of heads of households relative to college graduates of same year – i.e. 16 years of schooling. This ratio is interpreted as skill premium of x years of schooling relative to that of the average college graduate. Source: Table 37 in Kuhn and Ríos-Rull (2016). In Subfigure (c) growth rate dierences relative to counterfactual average growth rate are shown for year 2013 compared to 1989. The counterfactual average growth rate is computed by xing the distribution across education groups to 1989 levels and only changing wage rates of each education group according to the overall mean wage growth. Then wage rates of this ccounterfactual exercise are calculated relative to the actual wage rates of each educational group in 2013. Source: Table 35 in Kuhn and Ríos-Rull (2016).

These very unequal developments in wage rates and skill premia, have been unsurprisingly accompanied by increased inequality in income shares. According to the data of Piketty (2014), the income share of the top ten percent increased from 40 to over 50 percent between 1989 and 2013 in the United States. Using data from the Survey of Consumer Finances (SCF), Kuhn and Ríos-Rull (2016) report the income shares illustrated in Figure 2 for the United States in 2013. Employing the distributional jargon of Piketty (2014), the upper class – making up merely ten percent of the population – capture close to half of the income share. The poor, on the other hand, while just as populous, have a measly one percent share of national income. Even within the upper class inequality is skewed. The whole nine percent of the well-to-do own 27.3% of total

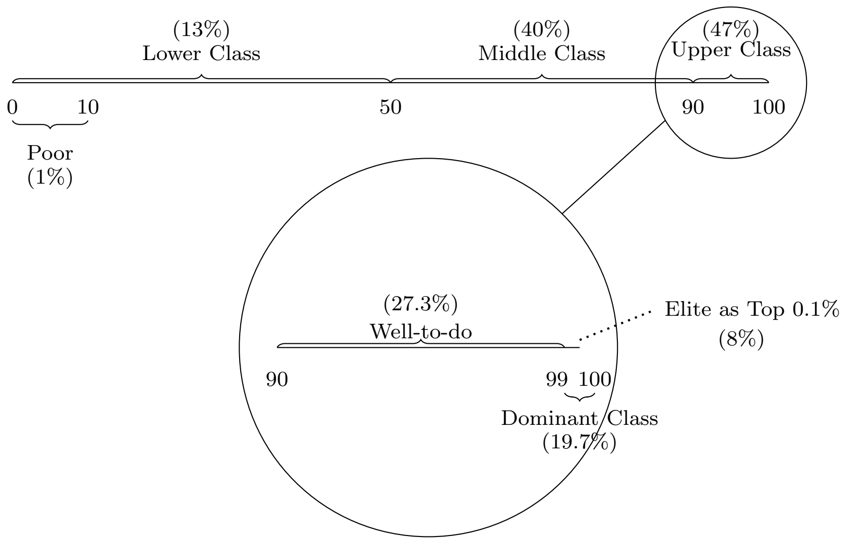

Figure 2: **Income Shares in the United States in 2013.**
The horizontal axis depicts the population share ordered by income, i.e. income percentiles. Segments of the population are categorized based on the distributional jargon of Piketty (2014). In parentheses the income shares are given for the United States in 2013. Source on income shares: Kuhn and Ríos-Rull (2016).

income, while the lonely percent of the dominant class own almost 20%. Kuhn and Ríos-Rull (2016) also report that the Gini-coefficient of the US income has risen from 55 percent in 1989 to 58 percent in 2013. All in all, the income distribution is very skewed at the cross-section and the last thirty years have exhibited increased inequality in the United states.

The framework developed in this paper shows how this increased and skewed inequality is compatible with ambiguity averse workers – exposed to rising uncertainty about the prospects of their skills – making learning decisions which yield increasing and convex wage returns to schooling. In short, in my framework, affluent workers are less concerned with forgone income due to ambiguous skill obsolescence investing more in education. Hence, they optimally accumulate a learning glut which pays large dividends during unexpected transitions that exhibit increased ambiguity leading to a skewed income distribution. My analysis thus yields three main conclusions. First, ambiguity of learning prospects and diminishing aversion towards it results in increasing and mainly convex pattern in skill premia as a function of education. Indeed, affluent workers tend to invest more in learning as their concern with forgone income due to skill obsolescence is progressively lower than their less credentialed counterparts – a phenomenon dubbed *type dependence* by Gabaix et al. (2016). Second, the high convexity of skill premia in education predicted by the calibration exercises suggest that aversion levels reported by experimental results (Ahn et al., 2014) are underestimated in relation to that of the macroeconomy. This discrepancy is akin to that between measures of risk aversion in the micro and macro literature (cf. e.g. Chetty (2006)) and can in turn be reconciled by observations on diminishing patterns in aversion (Baillon and Placido, 2019). It is, in other words, reasonable for individuals to be more averse when the stakes of their decisions concern lifetime income through educational attainment. Thirdly, unexpected transitional episodes with heightened ambiguity are compatible with increased inequality – a development documented during the last three decades (Piketty and Zucman, 2014; Kuhn and Ríos-Rull, 2016).

The rest of the paper is organized in the following manner. Section 2 identifies the related literature, embedding it in its scholarly context. Section 3 makes explicit the model environment – detailing the assumptions, definitions, dynamics and comparative statics. Next in Section 4, data sources and targets of the calibration exercises are discussed followed by the results. Thereafter I discuss the main insights from the theoretical analysis and numerical exercises in Section 5. Finally, I conclude with some remarks in Section 6.

2 Related Literature

This study is related to three sets of literature in economics. First, it concerns itself with the subject matter in the macroeconomic literature on human capital accumulation under economic turbulence. Second, it relates to the literature on skill obsolescence and labor-displacing events such as trade deals increasing import competition and routine-biased technological change. It also employs the computational methodology and analytical tools pertaining to the field of distributional analysis. Thereby, this study illustrates the crucial role of diminishing ambiguity aversion with respect to prospects of learning in explaining increasing and convex pattern of skill premia in education. Below I elaborate on the my study's connection to previous investigations in the mentioned fields.

First, 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). In particular, this paper is closely related to the subfield on the patterns of skill accumulation under economic turbulence. Ljungqvist and Sargent (1998, 2008) and Lalé (2018) define turbulence as the risk of instantaneous skill loss when a worker is exogenously separated from his job. Moreover, Bertola and Ichino (1995) and Ljungqvist and Sargent (1998) interpret the findings of Gottschalk and Moffitt (1994) on increased earnings instability as evidence for more turbulent times. Lalé (2018) replicates these results. Furthermore, Gottschalk and Moffitt (2009) provides an overview of evidence on rising economic turbulence. None of these studies, however, investigate the impact of ambiguity on human capital investment, the novel contribution of this current study. Empirical inquiries into the role of ambiguity in agents' learning decisions are few. Some recent notable work include Jensen (2010), Altonji, Blom and Meghir (2012) and Giustinelli and Pavoni (2017). Looking at secondary school and college students, a main insight of these studies is that individuals are averse to education that has ambiguous returns. Findings in this study are compatible with these dynamics.

Secondly, in terms of subject, this study overlaps with studies of skill obsolescence and labor-displacing economic changes. An overview of human capital obsolescence due to organizational and procedural issues is given in De Grip and Van Loo (2002). Aubert, Caroli and Roger (2006), moreover, provides evidence of age-dependent skill obsolescence by showing lower wage-bill of older workers in innovative firms. Furthermore, routine-biased technological change and job market polarization provide indirect evidence of some skills losing their monetized status (Autor, Levy and Murnane, 2003; Goos, Manning and Salomons, 2009, 2014). Trade deals paved the way for firms to offshore and outsource production or services far away and lower domestic wages through increased import competition (Bhagwati, Panagariya and Srinivasan, 2004; Ebenstein et al., 2014; Autor, Dorn and Hanson, 2016). Moreover, frameworks of labor-displacing technological change has been developed and tested through studying the effect of industrial robots on local labor markets with significant impact identified – especially among the low-skilled workers (Acemoglu and Restrepo, 2018, 2020; Autor and Salomons, 2018; Graetz and Michaels, 2018; Edin et al., 2019). While these papers focus on the change in demand for skills, this study focuses on how workers' alter their attainment, and subsequently, supply of skill in response to said demand shifts. As such this paper does not provide a competing narrative but another piece of the puzzle in understanding shifts following unexpected transitional episodes that exhibit increased ambiguity.

Finally, the distributional literature supplies this study with insights and computational tools. Some of the empirical targets, for instance, are derived from historical investigations such as Piketty and Zucman (2014), Saez and Zucman (2016) and Zucman (2019). More notably, however, this current study falls into the category of theoretical and computational distributional analysis. A main objective of this literature is to discern and quantify the contributors to power laws in economics and finance generally (Gabaix, 2009), for sizes of specific objects such as cities (Gabaix, 1999) or firms (Axtell, 2001), or for levels of wealth (Gabaix et al., 2016), income (Toda, 2012) or consumption (Toda and Walsh, 2015). Gabaix (2016) provides a good overview. Typically the source is explained in terms of some idiosyncratic shock, impacting the outcome of agents. Heterogeneity is then achieved through randomness in growth, level or size conditioned by agents' decisions. Similarly in this study, I model heterogeneity in income through idiosyncratic shocks to human capital stock. I consider a geometric Brownian motion with reset – the latter part of which is interpreted as skill obsolescence. This process has been previously considered by different distributional macroeconomic inquiries such as Toda (2012), Toda and Walsh (2015), Gabaix et al. (2016) and Nuño and Moll (2018). Toda (2012) and Toda and Walsh (2015) in particular identify a double-Pareto structure with power laws at both lower and higher ends of the distribution. Unlike the (single) Pareto distribution which only describes the top income shares, a double-Pareto structure enables tracking of inequality both at the top and at the bottom. Moreover, Ali Akbari and Fischer (2020) provide evidence for the superiority of this process in modeling income inequality.

In short, this study is related to the literatures on human capital accumulation under economic turbulence, distributional analysis and labor-displacing economic shifts. It utilizes the continuous-time extension of the smooth-ambiguity framework in Skiadas (2013) – based on original work by Klibanoff, Marinacci and Mukerji (2005, 2009) – under diminishing ambiguity aversion (Baillon and Placido, 2019). Distributional measures in aversion are borrowed from Ahn et al. (2014). My approach is closest to the distributional tools in Toda (2012) and uses the conceptual tools of Gabaix et al. (2016). More specifically, I provide a microfoundation of how individuals with high income growth can emerge – a phenomenon coined as *type dependence* by Gabaix et al. (2016). This is achieved by a macroeconomic application of the smooth-ambiguity framework – similar to the works of Ju and Miao (2012), Collard et al. (2018) and Miao, Wei and Zhou (2019). However, whereas said inquiries investigate the impact of ambiguity in financial markets, I apply the framework in the context of labor markets and workers' related educational decisions. As such – in terms of scope – the paper is closest to the macroeconomic investigations into heterogeneous human capital accumulation under economic turbulence, particularly Ljungqvist and Sargent (1998, 2008), Gottschalk and Moffitt (1994, 2009) and Lalé (2018). The subject of the paper, moreover, also takes a footing among the investigations into the impact of labor-displacing shifts such as Ebenstein et al. (2014), Autor, Dorn and Hanson (2016), Graetz and Michaels (2018) and Acemoglu and Restrepo (2020). In particular, the empirical findings of Edin et al. (2019) are very closely related. They document that workers at the bottom of their occupations initial earnings distributions suffered considerably larger earnings losses following adoption of labor-displacing technology. They account for these earnings losses by reduced time spent in employment, and increased time in unemployment and retraining. In this study, I show that given diminishing aversion to ambiguity, less affluent workers cannot afford foregone income preceding obsolescence, and hence fail to prepare for their skills potentially becoming obsolete. Thereby, this study focuses on how workers' alter their attainment, and subsequently, supply of skill in response to said demand shifts.

3 Model Environment

In this section, I make explicit the formal structure of the model environment. First, I discuss why the smooth framework of ambiguity aversion is particularly useful for explaining the pattern of skill premia. Second, I formally state the problem which employees face, and how they optimize their learning decisions. Alongside these results I discuss the dynamics of labor income. Distribution and aggregate stock of human capital and labor are derived next. In tandem, I also emphasize some comparative statics and deduce some distributional measures. Finally, the production technology is specified. Proofs for propositions and corollaries are stated in the appendix.

3.1 Choice of Ambiguity Framework

In attempting to explain the increasing and convex pattern of skill premia in education, I focus on two facets of uncertainty that employees face in their decisions to accumulate human capital: risk where employees know both the possible uncertain outcomes and their probabilities; and ambiguity when said probabilities are wholly or partially unknown.[1] In order to grapple with

[1] Previously, increasing and convex skill premia has been modeled through direct assumption of increasing and convex productivity. Furthermore, considerations regarding human capital investment have been typically left out.

ambiguity, employees are forced to rely on their beliefs influenced by their attitudes, an approach dubbed subjective expected utility (SEU) maximization (cf. e.g. Segal, 1987; Klibanoff, Marinacci and Mukerji, 2005; Maccheroni, Marinacci and Rustichini, 2006; Abdellaoui et al., 2011; Gul and Pesendorfer, 2014, 2020; Peysakhovich and Naecker, 2017).[2] Attitudes refers to those that pertain to an individual's belief structure, and is distinguished from her preferences. These attitudes are called *doxastic* as they relate to beliefs. Doxastic attitudes could then be overly pessimistic, optimistic or neutral – i.e. correct on average. Veering off of the course of neutrality is here referred to as *doxastic deviation*.

Scholars are still debating which conceptual framework lays the best grounds for formal modeling of ambiguity aversion. This literature has a long and vibrant tradition starting most likely with Knight (1921) while continuing to date. Nevertheless, unlike risk aversion, decision making under ambiguity has not been canonized. Hence there exists a whole host of theoretical toolboxes available for the practitioner among which to choose. Epstein (1999) provides a general account on modeling decision making under uncertainty. Ambiguity-focused approaches include (but are not limited to) rank-dependent utility (Segal, 1987), Choquet expected utility theory (Gilboa, 1987; Schmeidler, 1989), maxmin approach (Gilboa and Schmeidler, 1989), probability weighting (Prelec, 1998), α-maxmin approach (Ghirardato et al., 2004), smooth ambiguity aversion (Klibanoff, Marinacci and Mukerji, 2005), variational preferences (Maccheroni, Marinacci and Rustichini, 2006), expected uncertain utility theory (Gul and Pesendorfer, 2014) and calibrated uncertainty (Gul and Pesendorfer, 2020).

Notably, however, experimental and empirical investigations such as Halevy (2007), Baillon and Bleichrodt (2015) and Cubitt, Van De Kuilen and Mukerji (2020) have shown that no particular theoretical framework can alone consistently explain the dynamics of decision making under ambiguity. Indeed these studies generally tend to suggest that a successful framework in this regard should be non-probabilistic altogether, which has delayed canonization of any one framework. In this study the choice of framework has been therefore guided by practical considerations, until such non-probabilistic frameworks are developed.

Hence, I have chosen to employ the smooth ambiguity model of Klibanoff, Marinacci and Mukerji (2005, 2009), or more specifically, its continuous-time extension in Skiadas (2013) for three main reasons. First, it allows for distinguishing between impacts of ambiguity, aversion to it (preferences), and beliefs (doxastic attitudes). Second, the smooth ambiguity framework has been applied in macrobehavioral inquiries, and thus its dynamics are better understood in these settings. Indeed this framework has been used in a multiplicity of articles recently exploring various premium puzzles in the finance literature. Notably Ju and Miao (2012) have explored the role of smooth aversion towards ambiguity in investments to account for the risk-free pre-

Acemoglu and Restrepo (2018) for instance assume an exponential productivity schedule without human capital accumulation. The approach in this paper captures said pattern absent direct assumption. Instead, productivity is represented by the agents stock in human capital. Moreover, increasing and convex pattern of skill premium is a consequence of ambiguity and workers diminishing aversion to it.

[2] For a broad treatment of uncertainty see Epstein (1999).

mium puzzles while keeping ambiguity constant.[3] Gollier (2011) and Collard et al. (2018) utilize the framework in explaining equity premium puzzles. Moreover, it assists Miao, Wei and Zhou (2019) in treating the variance premium. Hence, this paper falls into the same tradition while applying the tools in a study of labor markets instead. Namely, I explore the role of ambiguity vis-à-vis the *skill* premium while allowing for diminishing patterns in aversion based on recent experimental findings of Baillon and Placido (2019). Thirdly, experimental results by Ahn et al. (2014) provide an empirical distribution of smooth ambiguity aversion, which can be utilized as a point of reference.

3.2 Workers

Workers $i \in [0,1]$ with rate of time preference $\rho > 0$, divide their attention budget of unity between learning a_{it} and labor $\ell_{it} = 1 - a_{it}$ at each point in time t – similar to the seminal framework of Ben-Porath (1967). Their human capital stock h_{it} develops in continuous time in accordance with a process that has three main parts: learning, risk and obsolescence. First, by devoting attention to learning, new stock with magnitude $g(a_{it})h_{it}$ is added to the worker's human capital, where $g(.)$ denotes the learning function with diminishing internal returns ($g' > 0$, $g'' < 0$). If no learning occurs, skills atrophy at the rate $\delta_h > 0$, i.e. $g(0) = -\delta_h$. Second, the skill stock is subject to idiosyncratic risk $\sigma h_{it} dB_i(t)$ where $\sigma > 0$ is the degree of volatility and $B_i(t)$ a Brownian motion. This risk is a catch-all component which represents all that creates variation in human capital despite identical investments in learning. They include, but are not limited to, health variations, personal issues, professional network variation, serendipity, etc. It is important to note that this variation is heteroskedastic; that is, workers face increased volatility the more human capital they have.

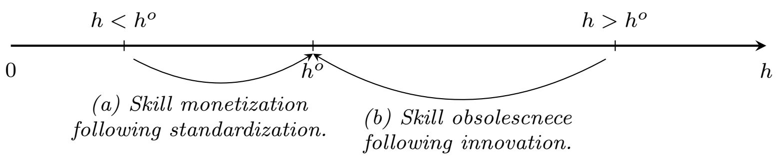

Figure 3: **Types of Skill Reset.**
The figure shows the two types of reset for human capital. (a) For workers with stock below reset level $h < h^o$, skills become monetized following standardization of production. (b) For workers with stock above reset level $h > h^o$, skills become obsolete following labor-displacing innovation.

Third, a worker's stock in human capital can become reset which is modeled by a Poisson arrival process $J_i(t)$. Upon obsolescence, skill stock is reset to a fixed level, i.e. $h_{it} = h^o \equiv 1$. As such, I follow the spirit of the definition on economic turbulence by Ljungqvist and Sargent (1998, 2008) and Lalé (2018) as the frequency of instantaneous skill loss. Consequently, there are two types of reset as illustrated in Figure 3. The conceptual framework extends that of Acemoglu, Gancia and Zilibotti (2012) to the dimension of human capital. Namely, Acemoglu, Gancia and Zilibotti (2012) identify standardization and innovation as the two engines of growth. Here we assume

[3]Ju and Miao (2012) use the recursive smooth ambiguity framework in Hayashi and Miao (2011) which, in turn, builds on the work of Klibanoff, Marinacci and Mukerji (2005, 2009).

that when certain production process is standardized, the associated skills become monetized. This event corresponds to the human capital stock resetting upwards for some workers with human capital level below reset level ($h < h^o$). Conversely, following labor-displacing innovations some skills become obsolete for workers with stock above reset level ($h > h^o$).

Hence, the full process of workers human capital accumulation is given by the following geometric Brownian motion with reset:

$$dh_{it} = g(a_{it})h_{it}dt + \sigma h_{it}dB_i(t) + dJ_i(t). \tag{1}$$

As mentioned earlier, skills become obsolete in accordance with a Poisson process – more specifically $J_i(t) \sim Poi(\lambda + \lambda^u u(h_{it}))$. Here, $\lambda > 0$ and $\lambda^u > 0$ are the frequencies of inherent obsolescence and long-term unemployment respectively. Moreover, $u(h)$ is the unemployment qualifier, and given by

$$u(h) = \begin{cases} \frac{1}{\ln h_{it}}, & h > h^o, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

In other words, the intensity of becoming indefinitely unemployed can be reduced by more specialization at the rate of $\ln h_{it}$. This can be thought of representing diminishing labor market tightness for workers with increased skill.[4]

The true frequency of inherent obsolescence $\lambda_{true} > 0$ is unknown to the worker, resulting in ambiguity. Thereby, workers are forced to make decisions based on their subjective beliefs μ. In order to keep the analysis salient, I assume a very simple belief structure. Namely, the worker considers two scenarios: a best case with obsolescence frequency $\lambda_s > 0$ and a worst case $\lambda_l > 0$. Intuitively, the best-case scenario posits lower frequency of skills becoming obsolete, i.e. $\lambda_s < \lambda_l$. Workers assign probability $p \in [0, 1]$ to best-case scenario and $1 - p$ to the worst case. The parameter p is then workers' degree of *optimism*. This is intuitive – higher optimism corresponds to larger p; that is, higher probability of the best-case scenario. Formally, the belief structure then follows a Bernoulli-type random variable $\Lambda : \Omega \to \{\lambda_s, \lambda_l\} \sim Be(p)$ where Ω is the state space and $p = \Pr(\Lambda = \lambda_s)$; that is,

$$\mu \sim \Lambda = \begin{cases} \lambda_s, & \text{with probability } p \in [0, 1], \\ \lambda_l, & \text{with probability } 1 - p. \end{cases}$$

Correspondingly, *ambiguity* is defined as $\xi := |\lambda_l - \lambda_s| = \lambda_l - \lambda_s$. Lower λ_s (less strenuous best-case scenario) or higher λ_l (more strenuous worst-case scenario) increases ambiguity *ceteris paribus*. This is intuitive as facing a larger span of worst and best cases, the agent is facing more severe uncertainty. Moreover, we define *neutral beliefs* as $p = \bar{p}$, where $\bar{p}$ is the level of optimism

[4]A result of this setting is that high-skill – and thus high-income – individuals are the ones not experiencing obsolescence for a long time in the stationary equilibrium. Since, however, obsolescence frequency is also affected by the level of human capital through $u(h)$, the mapping into high-income earners is well-motivated. The more human capital you have the less likely you are to have experienced skill obsolescence and to experience it again.

which yields the objectively true expected obsolescence frequency – i.e. $\bar{p} = \Pr(\bar{\Lambda} = \lambda_s)$ and $\mathbb{E}(\bar{\Lambda}) = \lambda_{true}$. *Doxastic deviation* is then intuitively defined as $p - \bar{p}$. When doxastic deviation is negative ($p - \bar{p} < 0$) the worker is overly pessimistic, and when it is positive she is overly optimistic.

Workers subsequently receive labor income at wage rate $w_t > 0$ per unit of human-capital-augmented labor supply $h_{it}\ell_{it}$. Hence, an individual i's total labor income at time t denoted by y_{it}^L is given by

$$y_{it}^L = w_t h_{it} \ell_{it}. \tag{3}$$

The workers objective is to maximize current-value of utility from consumption c_{it}. Workers cannot save ($c_{it} = y_{it}^L$). In other words, they are liquidity and credit constrained individuals who live on a hand-to-mouth basis. As such the only facet of saving that they can engage in, is their stock of human capital through learning. Instantaneous utility is set as logarithmic. In other words, workers have constant relative risk aversion of unity. As workers are ambiguity averse, their current-value is qualified by an uncertainty aggregator $\phi : \mathbb{R} \to \mathbb{R}, \phi' > 0, \phi'' < 0$ operating on the optimal current-value of utility conditional on the beliefs μ. More precisely, the employees face the smooth ambiguity problem (c.f. e.g. Klibanoff, Marinacci and Mukerji, 2005, 2009; Hayashi and Miao, 2011; Skiadas, 2013):

$$\max_{a_{it}} \mathbb{E}_\mu \, \phi \left(\mathbb{E}_t \int_t^\infty \ln(c_{is}) e^{-\rho(s-t)} ds \right). \tag{4}$$

A major advantage of the smooth ambiguity framework, is that it allows us to define workers' absolute ambiguity aversion $A(V)$ – similar to the risk-aversion counterpart – as

$$A(V) = -\frac{\phi''(V)}{\phi'(V)}. \tag{5}$$

where V is the current-value for a specific rate of obsolescence λ

$$V(h_{it}; \lambda, a_{it}^{**}) := \mathbb{E}_t \int_t^\infty \ln(c_{is}) e^{-\rho_i(s-t)} ds \bigg|_{a_{it}^{**}, \lambda}. \tag{6}$$

and a_{it}^{**} is the optimal choice of attention to learning for the worker's problem (4). The function ϕ models ambiguity aversion through displaying diminishing returns to aggregate current-value. Absolute ambiguity aversion $A(V)$ may be constant or diminishing in the population. Experimental results of Baillon and Placido (2019) show that absolute ambiguity aversion typically is decreasing. Theoretically I use the extension of the recursive smooth ambiguity approach of Klibanoff, Marinacci and Mukerji (2005, 2009) from discrete to continuous time in Skiadas (2013). This extension includes geometric Brownian processes with Poisson jumps. Importantly, Skiadas (2013) shows that there is no certainty equivalent non-ambiguous preference (pure risk aversion) representation that could capture dynamics of decision under mixed uncertainty – that is; diffusion (Brownian) and jumps (Poisson). This is a crucial result as it indicates that given skill obsolescence, we cannot replicate the dynamics under ambiguity aversion with modified patterns of risk aversion. In other words, ambiguity aversion produces unique predictions that risk aversion does not.

A worker's optimal labor supply is $\ell_{it}^{**} = 1 - a_{it}^{**}$. Thus, her labor income denoted by y_{it}^{L} is given by

$$y_{it}^{L} = w_t h_{it} \ell_{it}^{**}, \tag{7}$$

and its process given by

$$dy_{it}^{L} = \underbrace{w_t \ell_{it}^{**} dh_{it}}_{\text{Marginal income return to human capital}} - \underbrace{w_t h_{it} da_{it}^{**}}_{\text{Forgone income due to learning}}. \tag{8}$$

This representation of the labor income process is key to much of the analysis in this study, and provides crucial intuition for the interplay between learning decisions and ambiguity aversion. Indeed, ambiguity aversion – much like loss aversion – renders agents overweighting potentially diminished utility streams due to forgone opportunities. In this particular case, the second term in the income process would be weighted more relative to the first one compared to the situation when the agent would be ambiguity neutral. In other words, if the agent chooses to invest in more human capital, she forgoes income – which she would otherwise have earned through working – for the benefit of enhanced future income. Under ambiguity aversion, and given the possibility of skill obsolescence, the worker dreads not having the opportunity to benefit from the increased stock of human capital, why she overweights the diminished utility stream from forgone income. Put differently, she overweights the worst-case scenario. The degree to which she overweights said scenario, depends of course on both the perceived ambiguity – reflected through the belief structure μ – and aversion to said ambiguity – represented by the aggregator ϕ.

When there is no ambiguity and the true rate of obsolescence $\lambda = \lambda_{true}$ is known, optimal attention to learning – denoted by a_{it}^{*} – is crisp and defined by

$$a_{it}^{*} := \arg\max_{a_{it}} \mathbb{E}_t \int_t^{\infty} \ln(c_{is}) e^{-\rho(s-t)} ds. \tag{9}$$

Under the following regularity assumption:

$$\lim_{a \to 0} g'(a) > \rho + \lambda, \tag{10}$$

which guarantees interior solutions, the crisp optimal attention to learning a_{it}^{*} is given by the first-order condition

$$\frac{1}{1 - a_{it}^{*}} = \frac{g'(a_{it}^{*})}{\rho + \lambda}. \tag{11}$$

We can learn several important insights from this equation. First, the income and intertemporal substitution effects of consumption due to an increase in human capital cancel out each other. This is a result of relative risk-aversion being set to unity by the logarithmic felicity structure. Hence, when knowledge of the obsolescence rate is crisp we have constant optimal attention to learning $a_{it}^* = a^*$. As such, optimal attention to learning a_{it}^* neither depends on time, human capital stock nor wage rates. Second, the only parameters that affect learning decisions are rate of obsolescence λ and time preference ρ plus any parameters defining the learning function g. Indeed, optimal attention to learning is the same for all workers $a_{it}^* = a^*(\lambda, \rho; g(.))$. It is decreasing in rates of time preference and obsolescence. Moreover, an increase in marginal learning productivity g', increases optimal crisp choice of attention-to-learning a_{it}^*. I summarize this discussion in the following proposition and subsequent corollary.

Proposition 3.1. *For the unambiguous model described in (1), (7), (9) and (10) above, the agent's optimal choice of learning a_{it}^* is independent of level of human capital and wage levels while being decreasing in the discount rate and frequency (intensity) of human capital obsolescence:*

$$\frac{d}{dh_{it}} a_{it}^* = 0, \qquad \frac{d}{dw_t} a_{it}^* = 0, \qquad \frac{d}{d\rho} a_{it}^* < 0, \qquad \frac{d}{d\lambda} a_{it}^* < 0.$$

Beyond the discount rate and frequency of human capital obsolescence, the agent's optimal choice of learning depends on the characteristics of the learning function.

Corollary 3.1.1. *Higher overall marginal rate of learning (i.e. higher marginal efficiency in learning) - corresponding to larger $g'(a), \forall a \in [0,1]$ - leads to higher optimal attention to learning.*

It is important to note that some agents are better off upon reset. These include those whose trajectories lead to a human capital stock below obsolescence level h^o. This threshold is denoted by $\tilde{h}(a^*)$, and is given by discounted optimal learning net volatility and unemployment frequency:

$$\tilde{h}(a^*) \equiv \exp\left(\frac{1}{\rho}\left[\lambda^u + \frac{1}{2}\sigma^2 - g(a^*)\right]\right).$$

Observe that more investment in human capital (larger a^*), distances the individual from these sets of trajectories while higher volatility σ and unemployment frequency λ^u brings her closer to them. I summarize this discussion in the following corollary which states that for high enough stock in human capital, the agent is worse off upon obsolescence.

Corollary 3.1.2. *The optimal value function $V(h_{it}; \lambda, a_{it}^*)$ is decreasing in the frequency of human capital obsolescence if and only if*

$$h > \tilde{h}(a^*) \equiv \exp\left(\frac{1}{\rho}\left[\lambda^u + \frac{1}{2}\sigma^2 - g(a^*)\right]\right). \tag{12}$$

Having solved the optimization in absence of ambiguity, we can solve the ambiguous problem in (4) as:

$$\max_{a_{it}^*} \mathbb{E}_\mu \, \phi\left(V(h; \lambda, a_{it}^*(\lambda))\right).$$

From here on I will suppress the indexes i and t. Hence, the smooth ambiguity aversion problem becomes a choice between a continuum of lotteries $\mathcal{L}(\Lambda, a^*(\lambda))$ with expectation

$$\mathbb{E}_\mu \left(\mathcal{L}(\Lambda, a^*(\lambda^*))\right) = p\phi(V(h; \lambda_s, a^*(\lambda))) + (1-p)\phi(V(h; \lambda_l, a^*(\lambda))). \tag{13}$$

Observe that

$$\max_{a^*} \mathbb{E}_\mu \, \phi\left(V(h; \lambda, a^*(\lambda))\right) = \max_{a^*} \mathbb{E}_\mu \, \mathcal{L}(\Lambda, a^*(\lambda)), \tag{14}$$

and hence we define:

$$a^{**}(\lambda^*) = \arg\max_{a^*} \mathbb{E}_\mu \left(\mathcal{L}(\Lambda, a^*(\lambda))\right), \tag{15}$$

where λ^* is the agent's *revealed frequency of obsolescence*; that is, the one according to which the agent chooses to act. The optimal supply of attention to learning under ambiguity is then given by the first-order condition:

$$p\phi'(V_s)\frac{\partial}{\partial a}V_s + (1-p)\phi'(V_l)\frac{\partial}{\partial a}V_l = 0 \tag{16}$$

where $V_s \equiv V(h; \lambda_j, a^{**})$, $j = s, l$. Intuitively, the optimal choice of attention to learning under ambiguity will be between the crisp choices in best and worst-case scenarios, i.e.

$$a^{**}(\lambda^*) \in [a_l^*, a_s^*], \qquad \text{and} \qquad \lambda^* \in [\lambda_s, \lambda_l].$$

where $a_j^* \equiv a^*(\lambda_j), j = s, l$.

The impact of changes in belief structure on optimal attention to learning, depends chiefly on the workers' affluence. Of course, the more human capital a worker has the more affluent she is. When agents are affluent to the point of being worse off upon obsolescence – i.e. $h > \tilde{h}(a^{**})$ – then increases in pessimism $1-p$ typically leads to more labor supply. Indeed, affluent individuals lose a lot of potential labor income if their skills become obsolete, i.e. set to obsolescence level h^o. Thus, upon intensified pessimism, they decide to extract more dividends from their stock of human capital. In other words, the stakes of forgoing income today increases, worsening the prospects of learning. Correspondingly, a shift towards optimism prompts the affluent worker to invest further in learning, as the prospects of accumulating human capital enhances.

Similarly, such affluent workers will typically invest less in learning and instead supply more labor upon positive shifts in ambiguity ($\Delta\xi > 0$). Notice the qualifier *typically.* Indeed, a further condition is that the affluent worker is also averse enough towards ambiguity, i.e. $A(V) > \underline{A}$, for some floor of aversion $\underline{A} > 0$. If so, the worker views increased ambiguity as worsening the prospect of further learning. Hence, she will give up less labor income compared to the episode preceding heightened ambiguity. We summarize the discussion in the following proposition, which is one of the main results of this study.

Proposition 3.2. *Assume that $h > \tilde{h}(a^{**})$ and that the optimal amount of attention to learn-*

ing is an interior point, $a^{**} \in (a_l^*, a_s^*)$. Then higher optimism and lower ambiguity regarding incumbent skill leads to more human capital investment if the agent has high enough ambiguity aversion. More precisely

$$\text{(a) } \frac{\partial}{\partial p} a^{**} > 0, \qquad \text{and} \qquad \text{(b) } \frac{\partial}{\partial \xi} a^{**} < 0 \tag{17}$$

where the comparative statics (a) has only (12) as a sufficient condition, while (b) requires further that

$$A(V_s) > \underline{A} \equiv \frac{\frac{\partial^2}{\partial a \partial \lambda_s} V_s}{\frac{\partial}{\partial a} V_s \cdot \frac{\partial}{\partial \lambda_s} V_s}. \tag{18}$$

If $h \leq \tilde{h}(a^{**})$, then (17b) holds unconditionally.

These predictions are in line with the few empirical studies that exist on learning decisions under ambiguity. For instance, using survey data for eighth-grade boys in the Dominican Republic Jensen (2010) finds that the perceived returns to secondary school are extremely low - despite having high measured returns - leading to less investment in schooling. Similarly, Giustinelli and Pavoni (2017) - also using survey data - find that advancing Italian middle-schoolers avoid high-school subject tracks that are perceived to have ambiguous returns. Altonji, Blom and Meghir (2012) finds similar results for college students choice of career paths.

It is useful to formulate Proposition 3.2 for the agents' revealed frequency of obsolescence λ^*. Hence, we additionally state and prove the following corollary.

Corollary 3.2.1. *Assume that $h > \tilde{h}(a^{**}(\lambda^*))$ and that the optimal amount of attention to learning is an interior point, $a^{**} \in (a_l^*, a_s^*)$. Then higher optimism and lower ambiguity regarding incumbent skill leads the agent to act according to lower revealed frequency of obsolescence λ^*, if she has high enough ambiguity aversion. More precisely*

$$\text{(a) } \frac{\partial}{\partial p} \lambda^* < 0, \qquad \text{and} \qquad \text{(b) } \frac{\partial}{\partial \xi} \lambda^* > 0 \tag{19}$$

where the comparative statics (a) has only (12) as a sufficient condition, while (b) requires further that 18 holds. If $h \leq \tilde{h}(a^{**}(\lambda^*))$, then (19b) holds unconditionally.

Proof. This result is a direct consequence of Proposition 3.2 and the fact that $\frac{\partial}{\partial \lambda^*} a^{**} < 0$ by definition of revealed frequency of obsolescence λ^*. $\square$

3.3 Distribution of Human Capital

When the obsolescence frequency λ is ambiguous, the agents will act in accordance with their revealed level λ^*, while the economy is impacted by the true level λ_{true}. The distribution of human capital is thus specified as $\psi(h, t; a^{**}(\lambda^*), \lambda_{true})$. It is found by solving the Kolmogorov-

Forward equation

$$\frac{\partial}{\partial t}\psi = -\frac{\partial}{\partial h}\left(g(a^*)h\psi\right) + \frac{\partial^2}{\partial h^2}\left(\frac{\sigma^2 h^2}{2}\psi\right) + \left(\lambda + \lambda^u u(h)\right)\left[\psi(h^o, t) - \psi\right] \tag{20}$$

subject to $\displaystyle\int_t^\infty \int_0^\infty \psi(h, s; a^{**}(\lambda^*), \lambda_{true}) dh ds = 1.$

In order to derive the distribution, we first assume no ambiguity – that is $\lambda = \lambda_{true}$ is known – and no unemployment frequency $\lambda^u = 0$. Moreover, we will only solve the equation in the stationary equilibrium – $\frac{\partial}{\partial t}\psi = 0$ – or in other words, when the distribution no longer changes over time. Then, since optimal attention to learning a^* is independent of the stock of human capital h, the solution to equation (20) is a double Pareto distribution with mode at h^o:

$$\psi(h; a^*(\lambda), \lambda) = \begin{cases} Qh^{-\varsigma_-(\lambda)-1}, & \text{for} \quad 0 \leq h \leq h^o, \\ Qh^{-\varsigma_+(\lambda)-1}, & \text{for} \qquad h > h^o. \end{cases} \tag{21}$$

where $Q = \frac{\varsigma_+\varsigma_-}{\varsigma_+ - \varsigma_-}$ is a normalizing factor[5] plus $\varsigma_-(a^*(\lambda), \lambda)$ and $\varsigma_+(a^*(\lambda), \lambda)$ are

$$\varsigma_\pm(a^*(\lambda), \lambda) = \frac{1}{2}\left[1 - \frac{2g(a^*)}{\sigma^2} \pm \sqrt{\left(1 - \frac{2g(a^*)}{\sigma^2}\right)^2 + 8\frac{\lambda}{\sigma^2}}\right]. \tag{22}$$

We can now derive two popular measures of inequality for this distribution, namely, the Gini coefficient $\text{Gini}_\psi(a^*(\lambda), \lambda)$ and top z shares of labor income $s_\psi(z; a^*(\lambda), \lambda)$ as function of inherent frequency in human capital obsolescence λ.

Proposition 3.3 (Gini-Coefficient and Top Shares of the Income Distribution)**.** *Let $\lambda^u = 0$. If the frequency of human capital obsolescence λ is unambiguous, the Gini-coefficient $\text{Gini}_\psi(a^*(\lambda), \lambda)$ is equal to,*

$$\text{Gini}_\psi(a^*(\lambda), \lambda) = \frac{2\varsigma_+^2 - 2\varsigma_+\varsigma_- + 2\varsigma_-^2 - \varsigma_+ - \varsigma_-}{(\varsigma_+ - \varsigma_-)(2\varsigma_+ - 1)(1 - 2\varsigma_-)}, \tag{23}$$

and the top $z \in (\frac{\varsigma_+}{\varsigma_+ - \varsigma_-}, 1]$ shares of income $s_\psi(z; a^*(\lambda), \lambda) \in [0, 1]$ are given by,

$$s_\psi(z; a^*(\lambda), \lambda) = z^{1-1/\varsigma_+} \cdot \left(\frac{-\varsigma_-}{\varsigma_+ - \varsigma_-}\right)^{\frac{1}{\varsigma_+}} \cdot \frac{\varsigma_- - 1}{\varsigma_-}. \tag{24}$$

where $\varsigma_\pm \equiv \varsigma_\pm(a^*(\lambda), \lambda)$ is as in (22).

Larger upper-tail coefficient ς_+, fattens the tail, and thus typically is associated with larger inequality – i.e. enhanced Gini-coefficient and increasing top shares.

Finally, allowing for a positive unemployment frequency the distribution retains the double Pareto structure, while also adjusting powers $\varsigma_\pm(a^*(\lambda), \lambda)$ by the magnitude of human capital

[5] Q is such that $\int_0^\infty \psi(h; \lambda) dh = 1$.

logarithmically, i.e.

$$\varsigma_{\pm}(a^*(\lambda),\lambda) = \frac{1}{2}\left[1 - \frac{2g(a^*)}{\sigma^2} \pm \sqrt{\left(1 - \frac{2g(a^*)}{\sigma^2}\right)^2 + 8\frac{\lambda + \lambda^u u(h)}{\sigma^2}}\right]. \tag{25}$$

This will make the lower-end curvature less pronounced while increasing the fatness of the upper tail. In other words, as workers with more human capital face lower intensity in becoming indefinitely unemployed, they tend to have more time to accumulate capital, emphasizing their abundance relative to their counterparts with less stock.

Moreover, the average unemployment rate U is defined appropriately as

$$U \equiv \mathbb{E}_{\psi}[\lambda^u u(h)|_{(1+\varepsilon)h^o}^{\infty}] \tag{26}$$

where $\varepsilon > 0$ and is set to 40% in the calibrations. It can subsequently be calculated analytically as

$$\mathbb{E}_{\psi}[\lambda^u u(h)|_{(1+\varepsilon)h^o}^{\infty}] = \lambda^u \mathrm{Ei}(-\varsigma_+(1+\varepsilon)h^o)$$

where Ei(.) is the exponential integral. The rationale here for averaging over workers with human capital beyond $(1+\varepsilon)h^o$ is two-fold, one conceptual and one practical. First, averaging over values beyond the obsolescence level h^o is motivated by having defined unemployment only applicable to this group, as indicated by the unemployment qualifier $u(h)$ in (2). Conceptually, this construction corresponds to workers needing to exceed basic skill competency by some factor $1+\varepsilon$ in order to able to enter into the labor market. Second, the reason for introducing the threshold magnifier ε is due to the fact that the integral is mathematically ambiguous (undefined) over the whole support $[h^o,\infty)$. Its introduction, nevertheless, can conceptually be interpreted as saying that only workers with human capital level beyond the threshold $(1+\varepsilon)h^o$ find vacancy postings in line with their level of human capital frequently enough to be considered in the labor market.

3.4 Aggregate Stock of Human Capital

The distribution of human capital under ambiguity aversion will retain its previously derived structure while being modified in accordance with the optimal choice of learning a^{**} given the revealed rate of obsolescence λ^*, and the true – yet unobserved – rate λ_{true}. More precisely, it will have the double-Pareto structure in (21) with powers $\varsigma_{\pm}(h;a^{**}(\lambda^*),\lambda_{true})$ given in (25) (evaluated at a^{**}). Hence, in equilibrium the aggregate stock of human-capital-augmented labor supply HL is given by

$$HL = \int_0^{\infty} h(1-a^{**})\psi(h;a^{**}(\lambda^*),\lambda_{true})dh. \tag{27}$$

The comparative statics of the distribution $\psi(h; a^{**}(\lambda^*), \lambda_{true})$ with respect to optimism p and ambiguity ξ are, however, ambiguous. Hence, we cannot give a monotonous account of how changes in said variables will impact measures of inequality such as the Gini-coefficient and top shares of income derived in Proposition 3.3. Nevertheless, we can derive the following partial result.

Proposition 3.4 (Impact of Pessimism and Ambiguity on Inequality)**.** *When risk to human capital and workers' aversion to ambiguity are high enough, increased pessimism $1-p$ and ambiguity ξ will lead to fatter upper tail in the distribution of labor income $\psi(h; a^{**}(\lambda^*), \lambda_{true})$. More precisely, there exists a level of human capital risk $\tilde{\sigma} > 0$ such that when $\sigma > \tilde{\sigma}$, then*

$$\textit{(a)}\ \frac{\partial}{\partial p}\varsigma_+(h; a^{**}(\lambda^*), \lambda_{true}) > 0, \qquad \textit{and} \qquad \textit{(b)}\ \frac{\partial}{\partial \xi}\varsigma_+(h; a^{**}(\lambda^*), \lambda_{true}) < 0, \tag{28}$$

under conditions $h > \tilde{h}(a^{**})$ and (18).

Fatter upper tail typically indicates larger shares in labor income. Thus, the result above suggests that – given high enough risk to human capital – inequality deepens when employees become more pessimistic about the prospects of their careers and when said prospects become more ambiguous. Observe, however, that the impact on the measures of inequality is indeed nontrivial, as changes in optimism and ambiguity change not only the upper tail $\varsigma_+(h; a^{**}(\lambda^*), \lambda_{true})$, but also the lower tail $\varsigma_-(h; a^{**}(\lambda^*), \lambda_{true})$ and human-capital threshold of being negatively impacted upon obsolescence $\tilde{h}(a^{**})$. Nevertheless, as long as the impact on the upper tail is dominating, inequality will increase.

Observe that while wage rate w is still independent of workers' individual human-capital stock under ambiguity, labor supply ℓ^{**} no longer is. Hence, the distribution of labor income denoted by $\varphi(y^L; a^{**}(\lambda^*), \lambda_{true})$ will not necessarily be categorically equal to that of human capital, i.e. $\psi(h; a^{**}(\lambda^*), \lambda_{true})$ in (21). However, it will retain a qualitatively similar structure with slightly adjusted powers $\varsigma_\pm(h; a^{**}(\lambda^*), \lambda_{true})$. Aggregate labor income is then derived as $Y^L = wHL$.

An important note to be made is the fact that aggregate human-capital-augmented labor supply is typically less than the product of aggregate values of human capital and labor supply – i.e. $HL \leq H \cdot L$. Formally, this is simply due to the fact that the sum (or integral) of products is not equal to the product of sums (or integrals):

$$\begin{aligned} HL = &\int_0^\infty h\ell^{**}(h)\psi(h; a^{**}(\lambda^*), \lambda_{true})dh \\ &\qquad \leq \int_0^\infty h\psi(h; a^{**}(\lambda^*), \lambda_{true})dh \cdot \int_0^\infty \ell^{**}(h)\psi(h; a^{**}(\lambda^*), \lambda_{true})dh = H \cdot L. \end{aligned}$$

Intuitively speaking, since labor supply across the distribution depends on the stock of human capital that the workers possess, skills are augmented by less labor supply than in aggregate and vice versa. Put differently, not every individual's human capital is augmented with the whole economy's labor supply, and similarly not every individual's labor supply is augmented by the whole stock of human capital in the economy.[6]

3.5 Skill Premium Convexity in Education

Skill premium is defined as earnings of a unit supply of labor relative to a reference worker, for instance college graduates. More precisely, skill premium $\gamma(h; h_{ref})$ measures the wage rate of a worker with human capital stock h relative to some reference individual with h_{ref}:

$$\gamma(h; h_{ref}) \equiv \frac{wh}{wh_{ref}} = \frac{h}{h_{ref}}, \tag{29}$$

where the second equality follows from the constancy of the general wage rate w. The *skill premium curve* is defined as the graph obtained when plotting the the skill premium $\gamma(h; h_{ref})$ against attention to learning supplied given same level in human capital $a^{**}(h)$. An illustration is seen in Figure 4, where the bold line is the skill premium curve. The vertical and horizontal axes represent skill premia and attention to learning respectively.

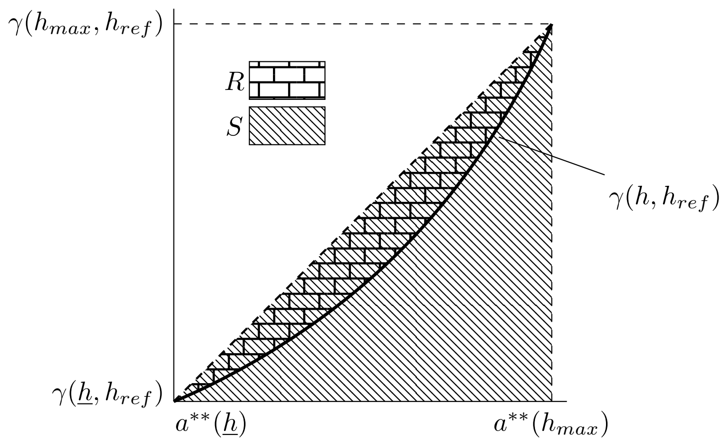

Figure 4: **Skill Premium Curve and Overall Convexity $\widetilde{\Gamma} = \frac{R}{S}$.**
The skill premium curve is $\gamma(h, h_{ref})$ plotted against optimal attention to labor $a^{**}(h)$ over the interval $[a^{**}(\underline{h}), a^{**}(h_{max})]$, where $\underline{h}$ is given by (32), and h_{max} is the maximum of human capital in the numerical grid in a calibration exercise.

If the improvement in skill premia due to an additional year of learning is more than the previous year, skill premia is said to be convex. Formally, skill premium at point h is convex in education if its marginal improvement per unit of attention-to-learning supplied is positive; that is,

$$\Gamma(h; h_{ref}) \equiv \frac{\partial^2 \gamma(h; h_{ref})}{\partial (a^{**}(h))^2} > 0. \tag{30}$$

[6]In order to capture high-skill workers also working more than the low skill, we would need to include disutility from labor in the utility structure of workers. In such scenario, workers divide their time between learning, labor and leisure, and it could potentially hold that high-skill agents not only continuously educate themselves more than the low-skill, but also work more than them. This would, in turn, intensify skill premium convexity and resulting income inequality.

Similarly, skill premia is concave in learning if the improvement is diminishing in learning – $\Gamma(h; h_{ref}) < 0$ – and linear if improvement is constant – $\Gamma(h; h_{ref}) = 0$. *Average skill premium convexity* is then given by

$$\mathbb{E}(\Gamma) = \int_{\underline{h}}^{\infty} \Gamma(h; h_{ref})\psi(h; a^{**}(\lambda^*), \lambda_{true})dh. \tag{31}$$

where $\underline{h} > 0$ is the threshold at which skill premia become increasing in education – formally,

$$\underline{h} \equiv \arg\inf \left\{ h \in \mathbb{R}_+ : \frac{\partial a^{**}}{\partial h} > 0 \right\}. \tag{32}$$

Since the value of average skill premium convexity $\mathbb{E}(\Gamma)$ spans over the whole spectrum of real numbers, it will make comparisons difficult. Hence, a normalized measure of skill convexity will be employed instead. Figure 4 provides some assistance in defining this measure, which is referred to as *overall skill premium convexity* and denoted by $\widetilde{\Gamma}$. Overall skill premium convexity is defined as:

$$\widetilde{\Gamma}(R, S) = \frac{R}{S}. \tag{33}$$

where $\widetilde{\Gamma} : \mathbb{R}_+ \times \mathbb{R}_+ \to [0, 1]$. The reader is invited to notice that this value has a similar relation to the skill premium curve, as the Gini-coefficient has to the Lorenz curve. Hence, just as the Gini-coefficient, when the skill premium curve is solely convex, overall skill premium convexity will be value between zero and one, easing comparisons among calibrations. Higher values of $\widetilde{\Gamma}$ entails stronger skill premium convexity, and vice versa.[7]

3.6 Capitalists, Production and Market Clearing

We assume that capitalists own the firms and thus are the drivers of physical capital accumulation in accordance with a simple neoclassical scheme. They save a constant portion $s \in (0, 1)$ of total output Y given by $F(K, HL)$ where F is a production function exhibiting constant returns to scale. Moreover, capital depreciates at constant rate $\delta_k > 0$. We define relative physical capital as $k = K/HL$. Then relative output of one final good consumed by workers is given by $y = Y/HL = F(k, 1) = f(k)$ by constant returns to scale. Subsequently, the process of relative capital accumulation is given by the deterministic process

$$\dot{k} = sf(k) - \delta_k k. \tag{34}$$

The number of owners of capital is very small relative to the continuum of workers. So their demand for final good is negligible. Hence, they do not impact the economy other than through their accumulation of physical capital. Assuming that F is such that f satisfies the Inada conditions[8] then there exists a *golden* level of relative capital $\tilde{k}$ at which relative capital growth

[7] When there is dominant concave segments, however, there $\widetilde{\Gamma}$ can become negative, down to a minimum of -1, as the skill premium curve will be north of the hypotenuse in Figure 4.

[8] They are $\lim_{k \to 0} f'(k) = \infty$ and $\lim_{k \to \infty} f'(k) = 0$.

is optimal and given by

$$sf(\tilde{k}) = \delta_k \tilde{k}. \tag{35}$$

Observe that while, at this level, relative capital $k(t)$ is constant, in absolute terms, however, physical capital $K(t)$ still grows at the same rate of human-capital-augmented labor supply $HL(t)$.

Finally, assuming risk-neutrality and time-indifference for owners of capital, the labor market clearing conditions are given by:

$$w = \vartheta \frac{\partial}{\partial HL} F(K, HL), \qquad \text{and} \qquad r = \frac{\partial}{\partial K} F(K, HL) - \delta_k,$$

determining rates of wage w and interest r respectively. Here $\vartheta \in (0, 1]$ is the portion of firms' revenue that is considered as point of departure in wage negotiations. This value is exogenously chosen in the calibration exercises. In reality, this portion depends partly on the unions' bargaining power and partly on interfirm competition (cf. e.g. Cahuc, Postel-Vinay and Robin, 2006). The more bargaining power the unions have and the more interfirm competition there exists, the less surplus firms can extract from the workers, and consequently $\vartheta \uparrow 1$. Formally, $\vartheta = 1$ represents the perfect market scenario with zero profits for the firm. This labor market clearing condition was suggested by McDonald and Solow (1981) and later microfounded by the seminal work of Mortensen and Pissarides (1994) in their search-and-matching model, versions of which are now standard in the literature that followed.

Equivalently, the clearing conditions are expressed as

$$w = (1 - \alpha)\vartheta \frac{Y}{HL}, \qquad \text{and} \qquad r = \alpha \frac{Y}{K} - \delta_k, \tag{36}$$

due to constant returns to scale in the production technology. Formally, the capital market equation for the interest rate r is not a clearing condition, but rather a constitutive one which determines the interest given the exogenous saving rate s. Larger saving leads to increased wealth-to-output ratio $\frac{K}{Y}$ in equilibrium, while higher worker revenue share ϑ enhances labor's share of income $\frac{Y^L}{Y}$.

4 Calibration

In this section I will produce some calibrations of the economy described in Section 3 above. First, I will specify the functional forms for the ones which were only partially determined. Second, I will present the data sources plus motivate the values employed in the calibrations in tandem with their empirical targets. Thirdly, I will illustrate the results for different regimes with constant and diminishing aversion to ambiguity – abbreviated as CAAA and DAAA respectively. I will show that the model under CAAA predicts a decreasing skill premium curve, which is contrary to empirics. I then will adopt DAAA under the zero-profit assumption, trying to match aversion quantiles in the experimental work of Ahn et al. (2014). The resulting exercise will exhibit exaggerated overall skill premium convexity. Two channels are then explored in matching empirical skill premia's convexity in education: increased ambiguity aversion and lower worker share of firm revenue. Allowing for heightened ambiguity aversion for lifetime outcomes in the macroeconomy, while well-motivated is seemingly inadequate in achieving this goal. Thus, lowering workers' share of firm revenue is deemed necessary in targeting empirical skill premia's convexity in education. Finally, I will perform a counterfactual analysis by inducing permanent ambiguity shifts to the benchmark calibration. I thus illustrate how transitional episodes exhibiting alleviated ambiguity allow affluent workers to extract dividends from their learning glut.

4.1 Functional Forms

In order to produce calibrations of the economy described in Section 3, we need to commit to specific structures for functions that were merely partially specified through their characteristics. These functions include the production technology $F(K, HL)$, learning technology $g(a)$ and current-value aggregator $\phi(V)$. For the production function, a standard Cobb-Douglas form is chosen

$$F(K, HL) = BK^{\alpha}(HL)^{1-\alpha},$$

where $B > 0$ an efficiency parameter interpreted as total factor productivity and $0 < \alpha < 1$ and $1 - \alpha$ the capital and labor shares in production respectively. The labor share in production $1 - \alpha$ equal to labor share of national income $\frac{Y_L}{Y}$ if firms' total revenue share is used in wage setting, i.e. $\vartheta = 1$.

For learning technology, I choose a functional form akin to Ben-Porath (1967) in the following form

$$g(a) = S\frac{a^{\theta}}{\theta} - \delta_h \tag{37}$$

where $S > 0$ is the self-productivity, $\theta \in (0, 1)$ the function's curvature and $\delta_h > 0$ rate of atrophy (forgetfulness). An identical functional form has been used in other related work on the macroeconomics of human capital accumulation such as Huggett, Ventura and Yaron (2006, 2011), Ludwig, Schelkle and Vogel (2012) and Guvenen, Kuruscu and Ozkan (2014).

The current-value aggregator $\phi(V)$ will determine workers' preferences regarding ambiguity. Since the theoretical results obtained in Section 3 are derived with conditions upon the absolute ambiguity aversion $A(V)$, a prudent functional choice would be that offers tractability for this value. Hence, we adopt two functional forms

$$\phi(V) = -e^{-\eta V} \tag{38}$$

(used also by Ahn et al. (2014)) and

$$\phi(V) = \beta V - e^{-\eta V}, \text{ where } \beta = \left(\frac{1}{\epsilon} - 1\right)\eta, \tag{39}$$

where $\epsilon \in (0, 1]$. The aggregator in (38) exhibits constant absolute ambiguity aversion (CAAA) $A(V) = \eta > 0$. The second aggregator in (39) exhibits decreasing absolute ambiguity aversion (DAAA) given by,

$$A(V) = \frac{\eta^2 \exp(-\eta V)}{\beta + \eta \exp(-\eta V)}, \tag{40}$$

which maps in accordance with $A : \mathbb{R} \mapsto (0, \eta)$. Diminishing pattern of aversion to ambiguity is reported in recent experimental findings of Baillon and Placido (2019). According to results in the current study, said diminishing aversion to ambiguity is key to seeing increased inequality and enhanced skill premium in the economy. Intuitively, for very well-off workers – i.e. with high current value V – the exponential term in (40) will be negligible, and hence such workers will act as if they were ambiguity neutral, i.e. $A(V) \downarrow 0$. Well-off workers of course are those with high stocks in human capital h. Analogously, agents with low human capital stock will exhibit higher absolute aversion to ambiguity with an upper bound of η.

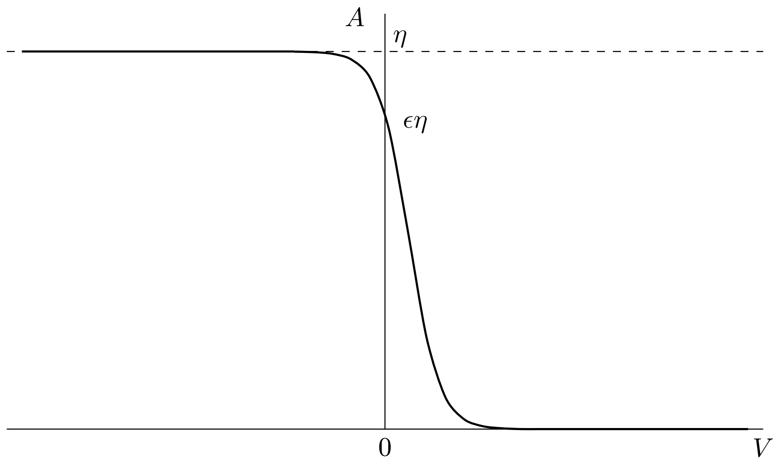

Figure 5: **Scheme of DAAA under aggregator $\phi(.)$ in (39).**

The vertical axis depicts absolute ambiguity aversion $A(V)$, while the horizontal axis represents the current value V. The aggregator $\phi(.)$ is exhaustively parametrized by η (upper bound on absolute ambiguity aversion) and ϵ (portion of ambiguity aversion at zero current-value). The aggregator $\phi(.)$ in (39) exhibits decreasing absolute ambiguity aversion given by (40).

The aversion scheme in (40) is depicted in Figure 5. Absolute ambiguity aversion A and current-value V are on the vertical and horizontal axises respectively. Under aggregator $\phi(.)$ in (39), ambiguity aversion does stretch monotonously decreasing from η to zero. Moreover, at current-value $V = 0$ the worker, will exhibit absolute ambiguity aversion equal to a portion $\epsilon \in (0, 1]$ of the maximum η seen in the population. For low values of $0 < \epsilon \lll 1$, zero current-value is achieved for some individual with very low human capital, typically below obsolescence level h^o. For $\epsilon = 1$, the DAAA aggregator in (39) collapses into the CAAA aggregator in (38).

4.2 Data, Targets and Calibration Values

The majority of the distributional measures of income is from the U.S. cross-sectional data in Survey of Consumer Finances (SCF) 2013 summarized in Kuhn and Ríos-Rull (2016). These include the Gini-coefficients, top-shares and quantiles in income.[9] The main target in this regard is a Gini-coefficient of 58%, which determines the choice of the volatility parameter σ. Important to note is that I have chosen to focus on the measures pertaining to income rather than earnings reported in Kuhn and Ríos-Rull (2016). Earnings tend to display higher inequality than income in the SCF data. Such targets can also easily be matched by slight increases in the volatility parameter σ. The earnings measure, however, includes only wages and salaries of all kinds, plus a fraction of business income – i.e. unambiguous labor earnings. As the goal of this study is to capture distributional dynamics pertaining to differences in stocks of human capital, such a limited scope of income streams are deemed to be too restrictive. Indeed there are other – possibly more ambiguous – income streams relating to human capital that are then excluded. These include, but are not limited to, residual business income, farm earnings, royalties, settlements, prizes, scholarships and grants. These latter streams of earning are included in the income measure, but not in earnings. Hence, targeting distributional measures based on income is deemed to be more prudent.

All measures on ambiguity aversion (average, median and quantiles) are based on the experimental results conducted by Ahn et al. (2014). I have chosen to focus on the full sample values reported. The targets have been mainly median and mean values of absolute ambiguity aversion – 0.02 and 0.21 respectively – which are achieved through adjustments done to the current-value aggregator ϕ. These adjustments are discussed in detail in the captions of corresponding tables. Finally, values on labor income share of 56.7% and unemployment rate of 7.4% in 2013 are reported by U.S. Bureau of Labor Statistics.

Some calibration values are directly derived from previous studies – a lot of which are standard in the literature – while others are chosen so as to match some targets, such as moments or averages in the data. Table 1 contains a complete list of these values pertaining to the benchmark calibration. Some of these values are sourced from references in the literature. Others have been adjusted to match moments in the data. Here follows a discussion on a select number of them starting with parameter choices for learning technology.

[9] Formally cumulative distribution function is,

$$\Phi(h; a^{**}(\lambda^*), \lambda_{true}) = \int_0^h \varphi(x; a^{**}(\lambda^*), \lambda_{true}) dx,$$

and subsequently the quantile of first s share in population is defined as

$$Q_m \equiv \Phi^{-1}(m; a^{**}(\lambda^*), \lambda_{true}),$$

where $m \in [0, 100]\%$ is some cumulative probability of interest. Corresponding definitions can be made for distributions of human capital and ambiguity aversion.

Table 1: **List of Benchmark Calibrated Values.**

Category	**Variable**	**Value**	**References, Matched Moments and Rationale**
Preferences and Beliefs			
Rate of time preference	ρ	0.05	Krueger, Mitman and Perri (2016).
Best case obsolescence frequency	λ_s	0	No obsolescence.
Worst case obsolescence frequency	λ_l	1	Normalization.
Optimism	p	0.1	$\mathbb{E}_\mu[\Lambda] = \lambda_{true}$.
Learning Technology			
Self-productivity	S	0.8	Cunha, Heckman and Schennach (2010).
Curvature of learning function	θ	0.75	Huggett, Ventura and Yaron (2006); Ludwig, Schelkle and Vogel (2012); Guvenen, Kuruscu and Ozkan (2014).
Depreciation of human capital	δ_h	0.015	Guvenen, Kuruscu and Ozkan (2014)
Obsolescence frequency	λ_{true}	0.9	Matching 8.6% annual job destruction rate (Michelacci and Lopez-Salido, 2007; Davis et al., 2010).
Unemployment frequency	λ^u	0.1	Matching average unemployment rate $U = 7.4\%$.
Human capital volatility	σ^2	1.1	Gini(Labor Income) = 58% (Kuhn and Ríos-Rull, 2016).
Production Technology			
Total factor productivity	B	0.498	Matching capital-to-output ratio $K/Y = 3.128$.
Capital share of production	α	0.36	Krueger, Mitman and Perri (2016).
Depreciation of physical capital	δ_k	0.028	Matching $r \in [8, 9]\%$ for total equity return (Saez and Zucman, 2016).
Saving rate	s	0.15	Chen, Karabarbounis and Neiman (2017); Saez and Zucman (2016).
Workers' revenue share	ϑ	0.768	Matching $\widetilde{\Gamma} = 34\%$.

I adopt the convention of no best-case obsolescence ($\lambda_s = 0$). The worst-case frequency will initially be set at unity ($\lambda_l = 1$), yielding normalized ambiguity at one ($\xi = 1$). The normalization makes the analysis comparable. Moreover, anchoring best-case obsolescence at zero frequency, removes the need to track $\tilde{h}$ for comparative statics, making the analysis more salient. In the benchmark calibration, optimism p is such that the agent has the correct expected beliefs regarding obsolescence – i.e. $p = \bar{p}$ and $\mathbb{E}_\mu[\Lambda] = \lambda_{true}$. Later, I will discuss the distributional impact of doxastic deviation. Finally, choices of parameters η and ϵ in the subjective aggregator ϕ will be explicated later. More precisely, η is chosen to match different upper percentiles of absolute ambiguity aversion reported in Ahn et al. (2014), while adjusting ϵ to match the median or mean of the distribution.

For learning technology three values are central: self-productivity S, curvature θ and obsolescence frequency λ_{true}. Self-productivity is derived from seminal empirical work in the psychometric literature on skill formation Cunha, Heckman and Schennach (2010). We use the second stage self-productivity in cognitive skill formation which is reported between 0.65 and 0.90. The second stage corresponds to individuals between ages 5-6 to 13-14. Relative to the first-stage levels in cognitive self-productivity, Cunha, Heckman and Schennach (2010) report an increasing trend in the second stage, suggesting that for the labor-productive years (say 15 to 75), this value should likely be set above one. However, other macroeconomic models of human capital accumulation report lower ability coefficients in estimates of the less granular – yet more comparable – seminal model of Ben-Porath (1967). For instance Huggett, Ventura and Yaron (2006, 2011) report mean ability of around 0.2 and 0.3 for curvature $\theta = 0.7$, while Ludwig, Schelkle and Vogel (2012) puts it at 0.16 for a curvature of $\theta = 0.65$. Hence, as the psychometric literature suggests self-productivity values larger than one while the macroeconomic studies on human capital accumulation points to values below 0.5, I have chosen the middle-ground at 0.8.

Regarding the curvature in learning, Huggett, Ventura and Yaron (2006) argues that it should be in the interval $\theta \in [0.5, 1)$. Ludwig, Schelkle and Vogel (2012), Huggett, Ventura and Yaron (2011) and Guvenen, Kuruscu and Ozkan (2014) use values 0.65, 0.7 and 0.8 respectively. The benchmark setting *ceteris paribus* is robust to this set of values and I chose, therefore, 0.75 as the benchmark. I follow the literature on economic turbulence in using job loss as a proxy for obsolescence (cf. e.g. Lalé, 2018). In particular, for determining the value of true obsolescence frequency λ_{true} we look at job destruction rates. Michelacci and Lopez-Salido (2007) use 3.2% and 6.8% for endogenous and exogenous quarterly job destruction to match the 10% separation rate into long-term unemployment reported in Den Haan, Ramey and Watson (2000) for the US job market. The annual rates are typically higher, for instance Davis et al. (2010) report 15% rate of job destruction for the entire US private sector over the period 1970-2005. Since not all destroyed jobs are due to obsolescence, we set $\lambda_{true} = 0.9$, interpreting $t = 1$ as a decade. This entails an annual obsolescence frequency of about 8.6%.[10] Indeed, job destruction rate serves as an upper bound for skill obsolescence. In other words, some jobs are destroyed due to forces of obsolescence and others through other forces (e.g. business cycle slumps). The calibration is robust to higher levels. On the other hand, too low levels in obsolescence frequency forces the density to collapse from a double Pareto into a plain Pareto distribution with bunching around origin. Other resulting moments, however, are less impacted.

The production technology contains several crucial parameters. Total factor productivity B is calibrated to match 68% of the 4.6 capital-output ratio (K/Y) reported in Piketty and Zucman (2014) for US in 2010. The rationale for matching only 68% of the ratio is the absence of precautionary saving among workers in the model, which accounts for at least 32% of wealth (Carroll and Samwick, 1998). The saving rate of 15% is the floor among top 10% of the wealth distribution in the US in 2010 (Saez and Zucman, 2016).[11] The rationale for focusing on the top 10% is the exogenous wealth accumulation process that is used in this study. This is prudent as the top 10% account for around 68% of the wealth accumulated (Saez and Zucman, 2016) – the same portion net precautionary motives (Carroll and Samwick, 1998). Capital depreciation rate $\delta_k = 2.8\%$ is calibrated to match total equity return of between 8% to 9% in US during 2006 according to Saez and Zucman (2016) which includes dividends and realized capital gains. Finally, the capital share of production is chosen at 36% similar to Krueger, Mitman and Perri

[10] Since obsolescence incidence is modeled as Poisson jump with intensity λ_{true}, time until obsolescence is exponentially distributed with expectation $1/\lambda_{true}$. The annual portion of individuals experiencing obsolescence is thusly given by $1 - \exp(-\lambda_{true} t) = 1 - \exp(-0.9 \cdot 0.1) \approx 8.6\%$.

[11] It also corresponds to the corporate saving rate according to Chen, Karabarbounis and Neiman (2017).

(2016).

4.3 Results

In this section, I will present results of the calibration. First, I will present the calibrated values of the production economy which will remain constant making comparable the different exercises. Second, I present calibration results under the assumption of constant absolute ambiguity aversion. Contrary to empirics, CAAA predicts diminishing skill premium. Hence, this exercise is followed by one calibrated to empirically verified assumption of DAAA (Baillon and Placido, 2019). Moments of ambiguity aversion are first matched to experimental evidence of Ahn et al. (2014) in a perfect market setting, showing exceedingly high overall skill premium convexity. Two channels are then explored in matching empirical skill premia's convexity in education: increased ambiguity aversion and lower worker share of firm revenue. Allowing for heightened ambiguity aversion for life-time outcomes in the macroeconomy, while well-motivated is seemingly inadequate in achieving this goal. Thus, lowering workers' share of firm revenue is deemed necessary in targeting empirical skill premia's convexity in education. Finally, I show how transitional episodes exhibiting alleviated ambiguity allow affluent workers to extract dividends from their learning glut.

4.3.1 Production Economy

Table 2 displays the predicted values of the production economy compared to their targets respectively. Due the exogenous nature of physical capital accumulation in the model, these values will remain constant as long as the calibrated production technology is not altered. Having put capital share in production to 36% ($\alpha = 0.36$) and a relatively low capital depreciation rate ($\delta_k = 2.8\%$), labor share of income is overestimated when firms gain zero profits ($\vartheta = 100\%$). The wealth-to-income ratio undershoots the target marginally, but remains reasonably still above 3. The interest rate only slightly exceeds the maximum target of 9%.

Table 2: **Predicted Values in the Production Economy**

			Model		
Category	**Variable**	**Target**	$\vartheta = 100\%$	$\vartheta = 87.5\%$	$\vartheta = 76.8\%$
Interest rate	r	$[8, 9]\%$	9.19%	8.88%	8.63%
Share of labor income	Y_L/Y	56.7	64%	55.99%	49.14%
Wealth-to-income ratio	K/Y	3.128	3.001	3.081	3.149

The table shows three measures of the production side of the economy, namely interest rate r, share of labor income Y_L/Y, and wealth-to-income ratio K/Y. Parameter values of the calibration are the ones mentioned in Table 1, except for worker shares of firm surplus ϑ which is reported for each column. For references on the targeted values see Table 1 and the associated discussion.

Decreasing revenue share in wage negotiations to 87.5%, the production economy displays behavior closer to the data. The interest rate is between 8 and 9% the interval reported in the data (Saez and Zucman, 2016). Labor's share of income is 56% very close to the target. However, the wealth-to-income ratio undershoots the target still. Moreover, the empirical worker shares of firm surplus ϑ tend to be lower. For instance Cahuc, Postel-Vinay and Robin (2006) report numbers mainly ranging between 20% to 70% depending on sector and skill seniority in the United

States.[12] Shimer (2005) reports the value of ϑ as being 72% for the aggregate economy. Hence, evidence suggests that we should decrease this value further. In the benchmark calibration, it is therefore set to 76.8% to match overall skill premium convexity $\widetilde{\Gamma}$ of around 34% in the data. While employing this value undershoots the labor share of income slightly, it improves on the predicted wealth-to-income ratio.

4.3.2 Constant Absolute Ambiguity Aversion

The predicted distributional values of economy under CAAA is displayed in Table 3. The fit of the model relative to data seems fairly good.[13] The best fit is achieved undoubtedly for the exercise with the lowest ambiguity aversion, equal to the full sample median (0.02) reported in Ahn et al. (2014). Income shares of the dominant class and the elite, the Gini coefficient and income ratio of upper to middle class have very little deviation from the data.

Table 3: **Predicted Measures under CAAA.**

Category	**Variable**	**Data**		**Model**	
			$\eta = 0.02$	$\eta = 0.21$	$\eta = 1.9$
Gini in human capital stock	Gini(H)	NA†	55.19%	50.96%	50.27%
Gini in income	Gini(Y_L)	58%	58.06%	54.34%	53.74%
Top 10% share in income	$s_{Y_L}(0.1)$	47%	50.03%	46.64%	45.25%
Top 1% share in income	$s_{Y_L}(0.01)$	19.68%	19.95%	17.10%	16.62%
Income ratio of upper to middle class	Q_{90}/Q_{50}	3.32	3.63	3.00	3.01
Income ratio of middle class to poor	Q_{50}/Q_{10}	3.46	2.64	2.68	2.69
Average unemployment rate	U	7.4%	3.15%	3.87%	4.02%
Average attention to learning	$\mathbb{E}(a)$	NA‡	45.22%	36.96%	35.69%
Average learning	$\mathbb{E}(g(a))$	NA§	57.27%	48.95%	47.62%

The predicted model values are derived under constant absolute ambiguity aversion. Moreover, $\epsilon = 1$, $\vartheta = 1$ and $\sigma^2 = 0.5$. Empirical inequality measures pertain to US SCF data in 2013 and are reported in Kuhn and Ríos-Rull (2016) tables 2, 3 and 45, plus figure 15. NA stands for *Not Available.* The values for the chosen levels of absolute ambiguity aversion are – in ascending order – equal to the median, mean and 95^{th} percentile reported in Ahn et al. (2014) table 3. †Human capital stock is a theoretical construct and intangible in general. Schooling years cannot be said to correctly represent this measure since learning is carried out significantly outside formal schooling years as well. ‡Average life-cycle attention to learning is not available in general. The closest comparable available variable is expected years of schooling divided by life-expectancy. According to Human Development Data in United Nation's Development Program, expected years of schooling in US was 16.1 and life expectancy 78.9 in 2013. Hence, portion of expected life-cycle schooling would be 20.41%. As expected, this value is lower than the values predicted by the model where attention to learning corresponds to all cognitive investments – both those within and those outside formal schooling years. §Average growth rate of human capital is unobservable.

However, a CAAA scheme fails to capture increasing and convex structure in skill premia, as indicated by Figure 6. It illustrates workers' optimal attention to learning under constant absolute ambiguity aversion expressed in (38) ($A(V) = \eta$ when $\epsilon = 1$). The values for the chosen levels of absolute ambiguity aversion are – in ascending order – equal to the median (0.02), mean (0.21) and 95^{th} percentile (1.9) reported in Ahn et al. (2014). Under CAAA, the more human

[12] In three out of sixteen skill-sector groups considered in Cahuc, Postel-Vinay and Robin (2006) the shares were outside this range with extremes at the lower (9% and 14%) and higher ends (95%).

[13] Indeed this is a consequence of the desirable properties of the geometric Brownian motion process with reset in distributional modeling of income (Ali Akbari and Fischer, 2020).

capital stock a worker has, the less she will invest in human capital. To understand this, recall the composition of labor income in (8). The more human capital the worker has, the more forgone income due to investment in learning it would entail. Moreover, workers remain equally concerned about an additional unit of utility stream forgone due to CAAA. These two channels create this diminishing pattern of attention to learning as a function of human capital stock. Consequently, skill premium is predicted to be decreasing in education under CAAA – which is contrary to empirical evidence displayed in Figure 1 and documented by the literature writ large (cf. e.g. Lemieux, 2006, 2008; Acemoglu and Autor, 2011; Autor, 2014).[14]

Another important caveat here is the external validity of using experimental results as targets for calibrations of the macroeconomy. For instance, there could be a discrepancy between measurements in ambiguity aversion at the micro and macro levels akin to the documented discrepancy in measures of risk aversion between the fields (cf. e.g. Chetty, 2006). Indeed, the microeconomic literature tends to report lower risk aversion in the population compared to studies based on macroeocnomic data. Quite reasonably the high levels of overall skill premium convexity above seem to suggest that absolute ambiguity aversion should be higher than the values reported in Ahn et al. (2014), at least in the macroeconomy.
Another concern is the empirically observed pattern of diminishing absolute ambiguity aversion in utility streams (Baillon and Placido, 2019). Taking such insights seriously, experiments using comparably low and finite compensations such as the ones performed by Ahn et al. (2014) would report lower aversion values, than if individuals were making decisions while facing significantly higher stakes. For most people, lifetime earnings impacted by their learning decisions are the highest stakes they ever face – indeed magnitudes larger than simple portfolio-selection exercises or betting on Ellsberg urn lotteries. The fact that the exercise with lowest level of ambiguity aversion in Table 3 produces the best fit is indicative of why an economy with decreasingly averse individuals would match the distributional evidence better. Indeed, such a scheme would allow for high levels of ambiguity aversion – which generally triggers low human capital investment – to be offset by the high magnitudes accumulated by less averse individuals.

Given the counterfactual result of diminishing skill premia in education and experimental evidence rejecting CAAA (Baillon and Placido, 2019), we consider calibrations with diminishing aversion to ambiguity. In line with empirical evidence, DAAA predicts an increasing and convex structure of skill premia in learning investments.

4.3.3 Diminishing Absolute Ambiguity Aversion

While constant aversions to risk and ambiguity are frequently assumed by practitioners, experimental results such as Baillon and Placido (2019) strongly suggest a diminishing pattern. Figure 7 illustrates workers' optimal attention to learning under diminishing absolute ambiguity aversion expressed in (39) ($0 < \epsilon < 1$) as a function of human capital stock. Individuals that are more well-off – i.e. possess larger stock of human capital – become less and less concerned with

[14]Also, in his seminal work, Mincer (1958) laid out conditions under which returns to training time could be convex.

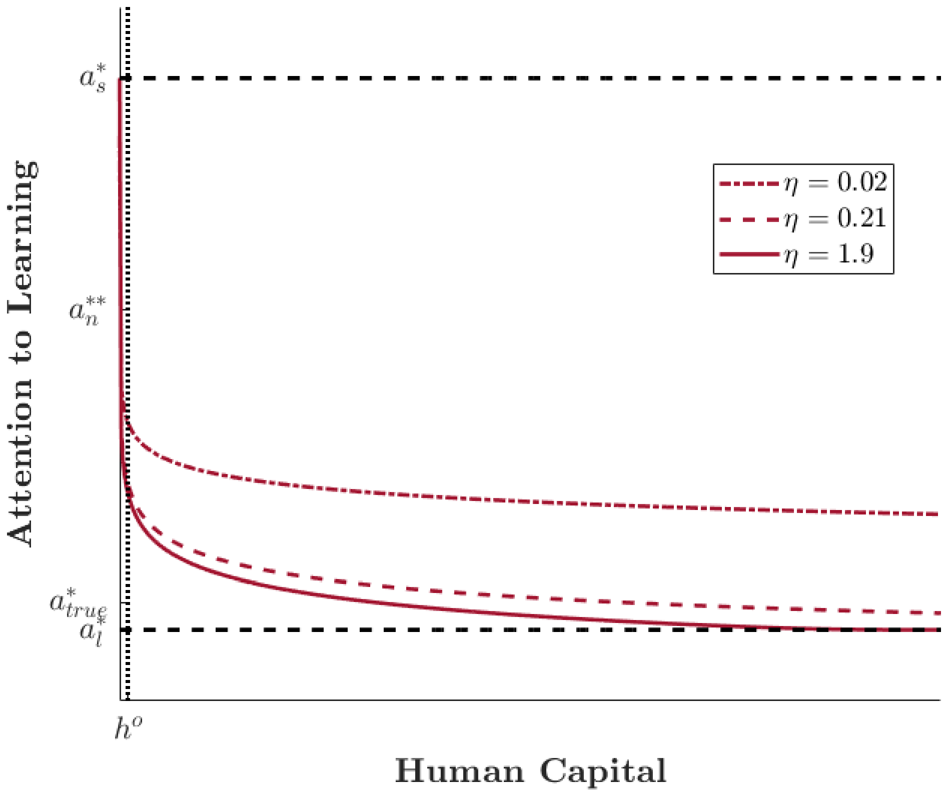

Figure 6: **Optimal Attention to Learning under CAAA.**

This figure illustrates workers' optimal attention to learning under constant absolute ambiguity aversion expressed in (38) as function of human capital h. Observe that $A(V) = \eta$ when $\epsilon = 1$. On the vertical axis optimal attention to learning a^{**} is reported, which spans between that in the worst-case a_l^* and best-case scenarios a_s^*. Also marked are optimal attention to learning under ambiguity neutrality a_n^{**} and optimal learning a_{true}^{**} if the agents knew the true obsolescence frequency λ_{true}. The horizontal axis depicts human capital stock h with a pin at the obsolescence level of human capital h^o. The values for the chosen levels of absolute ambiguity aversion η are – in ascending order – equal to the median (0.02), mean (0.21) and 95^{th} percentile (1.9) reported by Ahn et al. (2014). Under CAAA, the more human capital stock a worker has, the less she will invest in human capital.

their forgone income under this scheme. Hence, there exists a threshold $\underline{h}$ such that the marginal ambiguity-augmented current-value return of, on one hand, increased income due to learning is equal to the corresponding marginal loss due to forgone income. A formal characterization of $\underline{h}$ was given in (32). Thus, beyond $\underline{h}$ optimal investment in human capital increases.[15]

The location of this threshold depends on several variables and the interplay between them. Higher level of maximum aversion to ambiguity η tends to increase it, as does increased volatility σ. The intuition is easy to grasp. When η increases, the population becomes overall more

[15]Recall that the framework here abstracts from workers' financial savings. Including such saving as an instrument for insuring against loss of income due to skill obsolescence, would most likely reduce the incentive for human capital investment among the high-skill. Nevertheless, as long as learning efficiency is high enough, the pattern described should qualitatively remain the same.

averse to ambiguity, and an individual now must be even more well-endowed in human capital stock to be indifferent between marginal ambiguity-augmented current-value return and loss of increased stock of human capital and forgone labor income respectively. As it turns out, in most calibrations treated here, $\underline{h}$ is below obsolescence level of human capital h^o. This close-to-origin location of said threshold indicates that skill premium is increasing and convex for most of the population, except for a small fraction at the very lowest end of income.

Interestingly the dynamic here gives a granular account of type dependence which Gabaix et al. (2016) use to explain "superstar" agents with very high growth in human capital stock – and subsequently income. Indeed the analysis here suggests that under diminishing aversion to ambiguity, the more skilled – and consequently the more affluent – an agent is, the less she is concerned with forgone income. The reason is of course lower aversion of this individual towards ambiguity as reported by Baillon and Placido (2019). Hence, they engage in learning even further, increasing their skill premium, and likely alongside it, income inequality. However, as Figure 7 further displays, this level of investment in human capital is far beyond the optimum in absence of ambiguity a^*_{true}.
Below follows three sets of calibration exercises trying to match the pattern of the skill premium curve in 2013 displayed in Figure 1. First, I calibrate perfect market scenarios ($\vartheta = 1$), trying to match experimental moments of absolute ambiguity aversion in Baillon and Placido (2019). These sets of exercises are shown to predict too high convexity levels. I try to match empirical convexity through two measures: increased ambiguity aversion and lower firm revenue share for workers' wage-setting. Hence, the second set of calibrations allows for heightened aversion in lifetime outcomes while keeping the assumption that workers' wages are set based on total revenue of firms. These calibrations continue to overshoot empirical convexity levels however. Finally, I produce the third set of calibration exercises with workers' wages determined based on partial share of firm revenue and match empirical convexity of the skill premium curve.

Matching Experimental Aversion Moments in Perfect Markets

Table 4 summarizes the predicted distributional measures of economy under diminishing absolute ambiguity aversion when workers' wages are set based on total revenue of firms ($\vartheta = 1$). The table consists of two main categories measures pertaining to the distribution of the ambiguity aversion and those covering the economy writ large. The 95^{th} percentile of full sample absolute ambiguity aversion reported by Ahn et al. (2014) is around 1.9, why maximum aversion is put to equal to two ($\eta = 2$) in specifications (1) and (2). However, their results suggest that some agents have absolute aversion of up to 2.3. Hence, in order to accommodate such higher ceiling, specifications (3) to (5) are calibrated with maximum aversion set to four ($\eta = 4$). Adjusting ϵ allows us to target different statistics of the experimentally determined aversion distribution in Ahn et al. (2014). Specifications (1) and (3) target the median of full sample aversion, while (2) and (4) target the mean. The first and second couple of specifications respectively report aversion values lower and higher at other quantiles. Specification (5) is superior in overall fit to the aversion quantiles. It predicts lower aversion than what is reported by Ahn et al. (2014) on average, however.

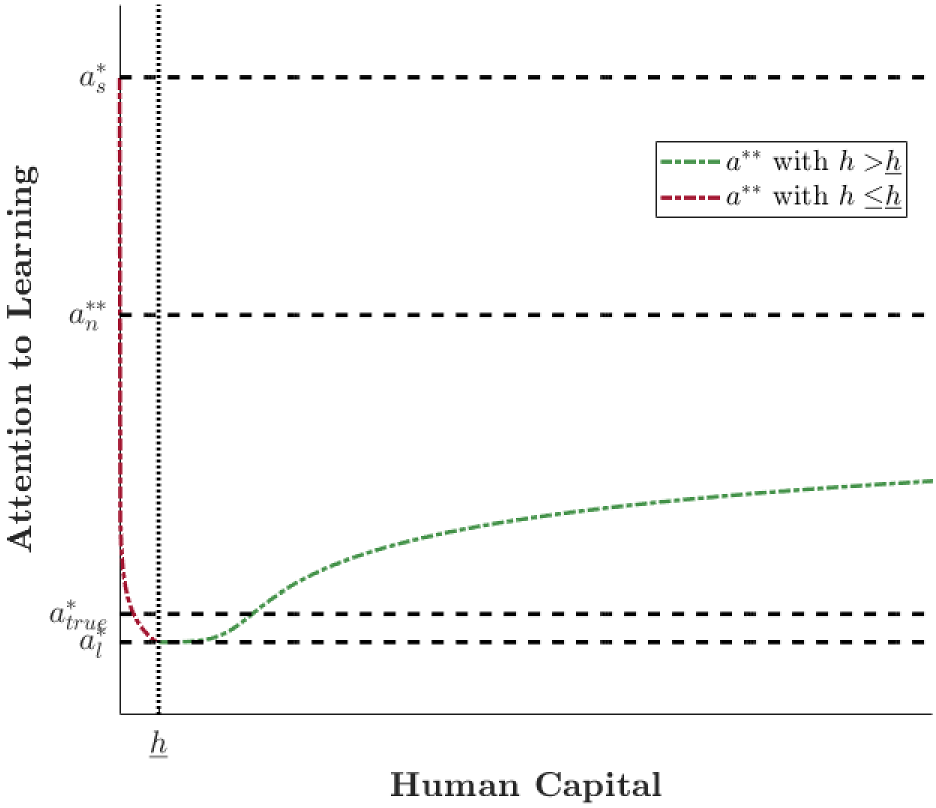

Figure 7: **Optimal Attention to Learning under DAAA.**
This figure illustrates workers' optimal attention to learning under constant absolute ambiguity aversion expressed in (38) as function of human capital h. Observe that $A(V) = \eta$ when $\epsilon = 1$. This is the specific result for the benchmark specification (15) in Table 6, which nevertheless illustrates the general dynamics very well. On the vertical axis optimal attention to learning a^{**} is reported, which spans between that in the worst-case a_l^* and best-case scenarios a_s^*. Also marked are optimal attention to learning under ambiguity neutrality a_n^{**} and optimal learning a_{true}^{**} if the agents knew the true obsolescence frequency λ_{true}. The horizontal axis depicts individuals' human capital stock h with a pin at the obsolescence level of human capital h^o plus the human capital threshold $\underline{h}$. A formal characterization of $\underline{h}$ was given in (32).

The distributional measures on income reported for the specifications calibrated continue to match the data well. Income Gini is around 58%. Quantile ratios of upper-to-middle classes are matched closely to values reported in the data. The income shares of the upper and dominant classes while close to 50% and 20% reported in the data, exceed the empirical levels consistently. Income share of the elite and quantile ratio of middle class to poor are underestimated, which is mainly a numerical construct.[16] Furthermore, unemployment is not very close to target.

[16] Numerically, I need to confine the calibration to a grid on an interval of human capital which was chosen as $[h_{min}, h_{max}] = [1\text{e-}6, 100]$. Extending the integrating support interval by increasing h_{max} allows us to better capture shares of income at rapidly more exclusive tops. The downside is loss of numerical accuracy at the smaller yet more populated interval surrounding obsolescence level of human capital h^o. A solution to this trade-off, is increasing the number of grid points, which makes the calibration taxing computationally, and thus, temporally.

Table 4: **Predicted Measures under DAAA**

Category	Variable	Data	Model (1) $\eta = 2$ $\epsilon = 3.35\text{e-}30$	(2) $\eta = 2$ $\epsilon = 3.8\text{e-}28$	(3) $\eta = 4$ $\epsilon = 7.2\text{e-}58$	(4) $\eta = 4$ $\epsilon = 3.8\text{e-}55$	(5) $\eta = 4$ $\epsilon = 3.8\text{e-}56$
Ambiguity Aversion							
Gini	Gini(A)	NA	51.62	29.68	64.18	43.86	49.00
Mean	$\mathbb{E}[A]$	0.21	0.03	0.21	0.05	0.21	0.13
Quantiles	$Q(A)$						
10 %	Q_{10}	0.00	0.00	0.09	0.00	0.04	0.02
25 %	Q_{25}	0.00	0.01	0.13	0.00	0.08	0.04
50 %	Q_{50}	0.02	0.02	0.20	0.02	0.17	0.09
75 %	Q_{75}	0.12	0.03	0.25	0.04	0.24	0.14
90 %	Q_{90}	0.50	0.05	0.33	0.09	0.37	0.24
Distributional Measures†							
Gini in human capital stock	Gini(H)	NA‡	60.65%	60.19%	60.69%	60.38%	60.57%
Gini in income	Gini(Y^L)	58%	60.18%	57.80%	59.96%	57.87%	58.75%
Top 10% share in income	$s_{Y^L}(0.1)$	47%	53.78%	51.64%	54.08%	51.95%	52.89%
Top 1% share in income	$s_{Y^L}(0.01)$	19.68%	21.49%	20.60%	21.44%	20.80%	21.10%
Top 0.1% share in income	$s_{Y^L}(0.001)$	8.07%	3.54%	3.51%	3.54%	3.53%	3.52%
Upper to middle class income	Q_{90}/Q_{50}	3.32	3.78	3.65	3.67	3.61	3.45
Middle class to poor income	Q_{50}/Q_{10}	3.46	1.96	2.10	2.30	2.09	2.27
Average unemployment rate	U	7.4%	2.08%	2.09%	2.08%	2.08%	2.08%
Average attention to learning	$\mathbb{E}(a)$	NA‡	60.63%	56.69%	60.31%	56.96%	58.43%
Average life-cycle learning	$\mathbb{E}(g(a))$	NA‡	71.79%	68.17%	71.50%	68.42%	69.78%
Skill Convexity	$\widetilde{\Gamma}$	34.04%§	99.92%	98.38%	99.89%	99.09%	99.53%

Ambiguity aversion measures are reported from full sample experimental results in table 3 of Ahn et al. (2014). Parameters are the same as in Table 1 except for $\sigma^2 = 0.5$ and $\vartheta = 1$. Specifications (1) and (3) target the median while (2) and (4) target the mean in empirical values of absolute ambiguity aversion. Specification (5) is the preferred baseline. NA stands for *Not Available*. †For information on references on data measures see caption of Table 3. ‡For a clarification and discussion see caption of Table 3. §As direct values on learning are unobservable, empirical measures of skill convexity become elusive. Skill convexity in terms of formal years of schooling, however, can be measured, but is not necessarily identical to the concept discussed here.

Finally, observe the high level of skill convexity reported for the calibrations. The values are high, around 99% just slightly short of the maximum. This result, while qualitatively consistent with the data (also cf. e.g. discussion of Acemoglu and Autor (2011) surrounding their figures 5 and 6, plus figure 6 in Autor (2014)), is quantitatively vexing. One conclusion is that the experimental aversion values reported in Ahn et al. (2014) are too low, which is consistent with implications of diminishing patterns in ambiguity aversion (Baillon and Placido, 2019). Indeed, comparatively smaller stakes which individuals face in experiments would then estimate lower aversion to ambiguity than when stakes are the lifetime earnings of individuals in the macroeconomy. This deviation in aversion to ambiguity is reminiscent of that of risk aversion between micro and macro literatures (cf. e.g. Chetty, 2006), which also can be reconciled using the insight of diminishing risk aversion provided by Baillon and Placido (2019).

The accuracy gains of such exercise was deemed small relative to the increased time cost. Similarly, reducing h_{min} increases accuracy for estimating quantiles at the lower end. However, putting h_{min} too low causes singularity issues of division by zero.

The increasing and convex structure of skill premium is displayed in Figure 8 for calibration specification (5) in Table 4. The horizontal axis represents portion of optimal attention to learning multiplied by number of productive years set to 45. The vertical axis exhibits the skill premium in logarithmic scale. The displayed convexity is excessive. Indeed, both years of learning and the overall convexity are overestimated, relative to the empirical evidence displayed in Figure 1. The first tenth order of premium is achieved by about three years of learning, the second by around an additional 1.5 years and the last by merely another six months. As indicated earlier, this high level of convexity suggests that the estimates of absolute ambiguity aversion in experimental literature such as Ahn et al. (2014) are lower than the levels which impact the macroeconomy. Indeed, this is a direct consequence of diminishing ambiguity aversion itself. Comparatively smaller stakes in experiments would then estimate lower aversion to ambiguity than when stakes are the lifetime earnings of individuals.

Next I will try to match overall convexity through two changes. First, I will allow for higher ambiguity levels relative to the experimental data. This will prove inadequate however. Second I will additionally allow for reduced worker share of firm revenue in wage setting ($\vartheta < 1$). Then we are able to match the convexity structure of the skill premium curve very well.

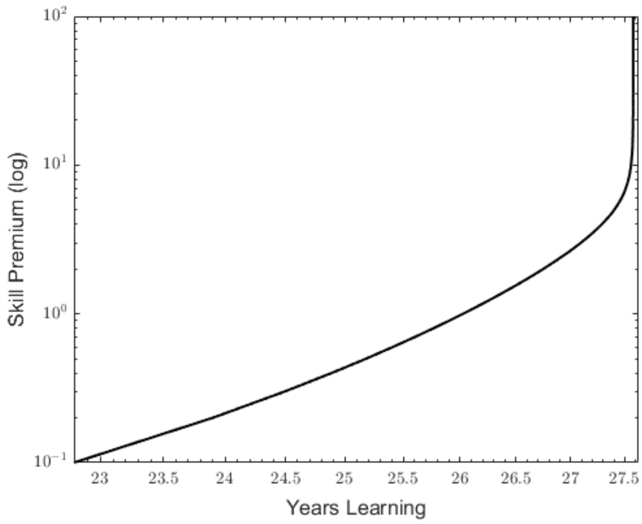

Figure 8: **Skill Premium Curve in Calibration Exercise (5).**

The horizontal axis depicts years spent learning out of an individual's first 45. The vertical axis is skill premium relative to obsolescnece level of human capital $\gamma(h; h^o)$. Both years of learning and the overall convexity are overestimated, relative to the empirics displayed in Figure 1.

Matching Convexity through Increased Aversion

A major conclusion from the previous set of calibrations was that experimental aversion values reported in Ahn et al. (2014) are too low, which is consistent with implications of diminishing patterns in ambiguity aversion (Baillon and Placido, 2019). Comparatively smaller stakes which individuals face in experiments would indeed record lower aversion to ambiguity than when stakes are the lifetime earnings of individuals. We also drew parallels between this deviation in aversion to ambiguity and the well-documented discrepancy in risk aversion measures among micro and macro literatures (cf. e.g. Chetty, 2006). Similarly the latter discrepancy can be reconciled using the insight of diminishing risk aversion provided by Baillon and Placido (2019).

Hence, we attempt to correct for this potential channel of positive bias in overall convexity by allowing for intensified ambiguity aversion. Table 5 shows the results of this set of calibration exercises. Specifications (6) to (10) incorporate increases in maximum absolute aversion to ambiguity η with magnitudes 50% to 75% relative to specification (5) in table 4. The parameter ϵ is adjusted to yield different spreads in aversion, which is witnessed by variation in average aversion levels. We also need to increase volatility σ^2 from 0.5 to 0.985 to match the empirical Gini coefficient of 58%.

The increase in ambiguity aversion indeed generates a lot of improvement in matching the empirical evidence. Comparing calibration exercises in Tables 4 and 5, we see that the latter yields lower top shares, which are closer to the data. Moreover, the quantile ratios, especially at the bottom of the distribution improve a lot. Moreover, the unemployment rate increases and is closer to the empirical level.

Table 5: **Predicted Measures under DAAA with Heightened Aversion.**

Category	Variable	Data	Model (6) $\eta = 6$ $\epsilon = 3.8e\text{-}67$	(7) $\eta = 6.25$ $\epsilon = 3.5e\text{-}69$	(8) $\eta = 6.5$ $\epsilon = 3.5e\text{-}72$	(9) $\eta = 6.75$ $\epsilon = 1.2e\text{-}75$	(10) $\eta = 7$ $\epsilon = 1.2e\text{-}80$
Ambiguity Aversion							
Gini	Gini(A)	NA	52.51	50.09	52.91	57.93	64.81
Mean	$\mathbb{E}[A]$	0.21	1.27	2.04	1.95	1.67	0.87
Distributional Measures†							
Gini in income	Gini(Y^L)	58%	58.51%	58.15%	58.26%	58.53%	59.39%
Top 10% share in income	$s_{Y^L}(0.1)$	47%	49.62%	49.35%	49.66%	49.77%	50.72%
Top 1% share in income	$s_{Y^L}(0.01)$	19.68%	19.40%	19.18%	19.21%	19.40%	19.81%
Top 0.1% share in income	$s_{Y^L}(0.001)$	8.07%	3.83%	3.81%	3.83%	3.81%	3.75%
Upper to middle class income	Q_{90}/Q_{50}	3.32	3.17	3.17	3.06	3.17	3.38
50-30 ratios in income	Q_{50}/Q_{30}	1.64	1.44	1.44	1.44	1.44	1.44
Middle class to poor income	Q_{50}/Q_{10}	3.46	3.88	3.79	3.96	3.86	3.87
Mean-to-Median ratio	$\mathbb{E}(y^L)/Q_{50}$	1.85	1.86	1.81	1.83	1.87	1.98
Average unemployment rate	U	7.4%	5.33%	5.36%	5.35%	5.32%	5.23%
Skill Convexity	$\widetilde{\Gamma}$	34.04%§	78.62%	77.69%	77.42%	86.20%	89.20%

Parameters are the same as in Table 1 except for $\sigma^2 = 0.985$ and $\vartheta = 1$. Specifications (6) to (10) incorporate increases in maximum absolute aversion to ambiguity η with magnitudes 50% to 75% relative to specification (5) in table 4. NA stands for *Not Available.* †For information on references on data measures see caption of Table 3. ‡For a clarification and discussion see caption of Table 3. §As direct values on learning are unobservable, empirical measures of skill convexity become elusive. Skill convexity in terms of formal years of schooling, however, can be measured, but is not necessarily identical to the concept discussed here.

More importantly, the chief purpose of these exercises, namely lowering overall skill premium convexity is achieved. It decreases by around 10% to 15%. Nevertheless, the overall convexity

$\widetilde{\Gamma}$ is still more than twice the empirical level of 34% observed. Hence, it seems as though this correction is insufficient by itself. Another observation that strengthens this conclusion, is the fact that there is no clear correlation between different increases in maximum ambiguity aversion η or variations average ambiguity $\mathbb{E}(A)$ on one hand, and variations in overall skill convexity $\widetilde{\Gamma}$ on the other. If any increase in either of η or $\mathbb{E}(A)$ yielded monotonous decreases in $\widetilde{\Gamma}$, we would have had a path to match empirical levels of skill premia's convexity in educational attainment.

Hence, we move on to the second measure for matching skill premium convexity, namely decreasing share of firm revenue in wage setting ($\vartheta < 1$). However, we will keep the increased ambiguity aversion for two reasons. First, as discussed, higher aversion levels towards ambiguity is documented through experiments (Baillon and Placido, 2019) and thus relevant for lifetime outcomes. Second, as we saw in Table 5, overall match of the income distribution is improved. We choose therefore the aversion parameters in Specification (9) – which has arguably the best overall fit – as a baseline for the next set of calibration exercises.

Matching Convexity through Lower Worker Revenue Share

When firm revenue share serving as basis for wage determination is reduced, unsurprisingly the aggregate wage level w decreases. Consequently, labor income of an individual worker y^L diminishes and so does its aggregate share as part of GDP $\frac{Y^L}{Y}$. Recall that wage rate enters labor income multiplicatively, i.e. $y^L = wh\ell$. As such, it disproportionately affects high earners, and diminishes the "superstar" effect. In other words, it scales down the skill premium. Gabaix et al. (2016) aptly dub this multiplicative impact *scale dependence*, and recognize its significance for the degree of convexity in income profiles.

Table 6: **Predicted Measures under DAAA Adjusting for Worker Share of Surplus.**

Category	**Variable**	**Data**	**Model** (11) $\vartheta = 100\%$	(12) $\vartheta = 87.5\%$	(13) $\vartheta = 85\%$	(14) $\vartheta = 80\%$	(15) $\vartheta = 76.8\%$
Distributional Measures†							
Gini in income	$\text{Gini}(Y^L)$	58%	60.22%	57.46%	57.55%	57.99%	58.30%
Top 10% share in income	$s_{Y^L}(0.1)$	47%	51.34%	46.01%	45.22%	46.29%	46.65%
Top 1% share in income	$s_{Y^L}(0.01)$	19.68%	19.97%	17.69%	17.10%	16.36%	16.84%
Top 0.1% share in income	$s_{Y^L}(0.001)$	8.07%	3.85%	4.14%	4.18%	4.24%	4.33%
Upper to middle class income	Q_{90}/Q_{50}	3.32	3.18	2.82	3.00	3.04	3.04
50-30 ratios in income	Q_{50}/Q_{30}	1.64	1.70	1.64	2.08	2.08	2.08
Mean-to-Median ratio	$\mathbb{E}(y^L)/Q_{50}$	1.85	1.88	1.63	1.64	1.65	1.66
Average unemployment rate	U	7.4%	5.90%	6.32%	6.45%	6.81%	7.12%
Skill Convexity	$\widetilde{\Gamma}$	34.04%§	80.03%	71.36%	68.65%	56.97%	34.98%
Aggregate Measures							
Labor share of income	Y^L/Y	56.7%	64%	55.99%	54.39%	51.19%	49.14%
Wealth-to-income ratio	K/Y	3.128	3.001	3.081	3.097	3.129	3.149
Interest rate	r	[8, 9]%	9.2%	8.88%	8.82%	8.71%	8.63%

Parameters are the same as in Table 1. Specifications (11) to (15) have η and ϵ as specification (9) in Table 5. NA stands for *Not Available*. †For information on references on data measures see caption of Table 3. ‡For a clarification and discussion see caption of Table 3. §As direct values on learning are unobservable, empirical measures of skill convexity become elusive. Skill convexity in terms of formal years of schooling, however, can be measured, but is not necessarily identical to the concept discussed here.

Table 6 summarizes the results following this modification. We use the aversion parameters for specification (9) in Table 5, and increase the volatility from 0.985 to 1.1 in order to sustain the 58% target in income Gini. As expected, lowering the firm revenue share for wage determination ϑ is very efficient in decreasing the overall skill premium convexity, and subsequently matching the empirical value. A decline of revenue share in wage negotiations to 87.5% in specification (12), improves several aggregate measures relative to data. The interest rate is between 8 and 9% the interval reported for high revenue capital in Saez and Zucman (2016). Labor's share of income is 56% very close to the target. However, the wealth-to-income ratio continues to undershoot the target. Moreover, recorded worker shares of firm revenue in wage negotiations ϑ tend to be lower. For instance, for 13 out of 16 sector-skill categories, Cahuc, Postel-Vinay and Robin (2006) report numbers ranging between 20% to 70% in the United States. Moreover, Shimer (2005) reports the value of ϑ as being 72% for the aggregate economy. Hence, evidence suggests that we should decrease this value further. In the benchmark calibration (15), it is therefore set to 76.8% to match overall skill premium convexity $\widetilde{\Gamma}$ of around 34% in the data. While employing this value undershoots the labor share of income slightly, it improves on the predicted wealth-to-income ratio.

The distributional moments also mainly improve, especially for top shares. The shares of upper and dominant classes are no longer overestimated and are close to the empirical targets. Unemployment rate increases and approaches the target. The only point of slight deterioration is an overestimation of the median income, which manifests itself in the three ratios that include it. The departure is negligible nonetheless, especially in the light of matching overall skill premium convexity $\widetilde{\Gamma}$.

A match in overall skill premium convexity $\widetilde{\Gamma}$ does not guarantee a match in profile of the corresponding curve $\gamma(h, h_{ref})$, however. This is similar to the relation of the Gini-coefficient to income distributions. Namely, an infinite number of distributions can exhibit the same Gini-coefficient. In other words, the overall inequality documented by the Gini index can be due to inequality at the top, bottom or between sporadic clusters. In the same way, overall skill premium convexity $\widetilde{\Gamma}$ can be based on very different skill premium profiles $\gamma(h, h_{ref})$.
Hence, we need to check how the skill premium curve fares relative to data. As optimal attention to learning a^{**} is specified as a fraction, we need to fix the maximum number of years spent in learning. Figure 9 – with subfigures (a), (b) and (c) – shows the skill premium curve $\gamma(h, h_{ref})$ for specification (15) when this maximum set to 18, 19 and 20 years respectively. The reference point is college graduates $h_{ref} = h_{college}$ – i.e. 16 years of education. As is apparent, the match is exceptionally close, and a stark improvement compared to Figure 8. Interestingly the skill premium profile captures the slight concavity at the bottom. This pattern is particularly visible in subfigures (b) and (c).

Capturing the skill premium curve for higher number of years spent learning – while matching the empirical pattern displayed for the interval of 10 to 18 years – would most likely require allowing for higher overall skill premium convexity $\widetilde{\Gamma}$. Namely, the skill premium curve would be more convex beyond the recorded maximum of 18. We already see this pattern emerge in

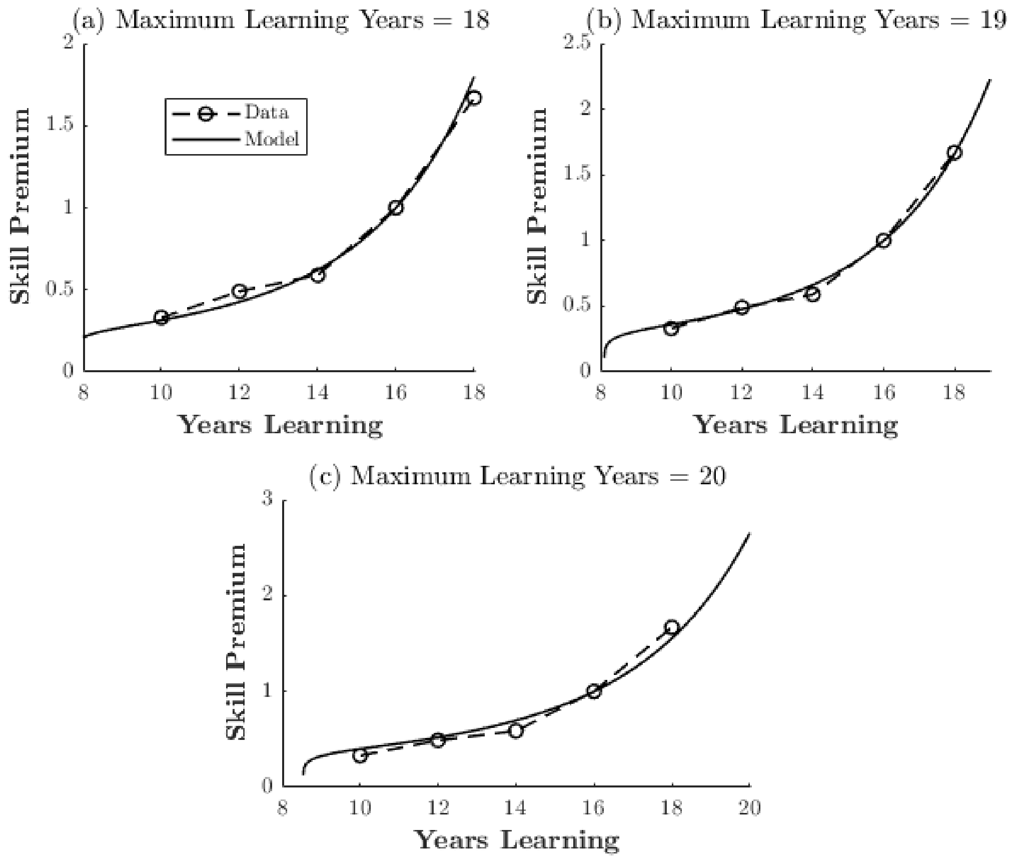

Figure 9: **Skill Premium Curve Relative to College Graduates $\gamma(h; h_{college})$.**
The results correspond to for benchamrk calibration (15). The values use in the calibration and all the targetted values are listed in Table 1. Subfigures (a), (b) and (c) illustrate the curve for maximum learning years set to 18, 19 and 20 years respectively. The empirical curve is that of the target year 2013 in the United States. The target is overall skill premium convexity $\widetilde{\Gamma} = 34\%$. In other words, the skill premium curve is not directly targeted.

subfigure (c) where skill premia are underestimated and overestimated at values before and beyond the reference level respectively – if only slightly so. Indeed, it is for this very reason that specification (15) provides a better matching in subfigure (b) when maximum education years is set at 19, slightly beyond the empirical threshold 18. Namely, the overall skill premium convexity $\widetilde{\Gamma}$ for this specification at almost 35%, slightly exceeds the empirical value at 34%.

Nevertheless, the very close match of the skill premium curve is remarkable for three reasons. First, the curve as a whole was not directly targeted but only the overall skill premium convexity $\widetilde{\Gamma}$. Second, the model is tractable – namely the functional choices are standard and physical capital accumulation is exogenous. Thirdly, it simultaneously matches different moments of the income distribution, including the well-sought-after top shares.

Now that we have identified a benchmark calibration in specification (15), and matched the empirical skill premium curve and the distribution writ large, we can move on to assess the consequences of shifts in ambiguity.

4.3.4 Impact of Ambiguity Shifts

We can now explore the impact of ambiguity shifts. The results in this section are thus numerical comparative statics resulting from relative changes in ambiguity $\frac{\Delta\xi}{\xi}$ relative to benchmark *ceteris paribus*. In particular, I vary ambiguity relative to benchmark, while keeping everything else constant, including expectations – i.e. $\mathbb{E}_{\mu}[\Lambda] = \lambda_{true}$ or equivalently $p = \bar{p}$. In other words, workers keep their beliefs neutral. If this is not satisfied, workers will instead exhibit doxastic deviation, and the comparison will not be prudent.[17] It is important to note that the insights of this exercise, however, should not be interpreted as a quantification of the contribution of ambiguity shifts to increased skill premium convexity during the recent decades. Measuring the significance of ambiguity shifts would require survey data on perceived uncertainty and goes beyond the scope of this study.

Recall that given diminishing aversion to ambiguity, affluent workers tend to invest more in learning as their concern with forgone income due to skill obsolescence is progressively lower than their less credentialed counterparts. The affluent workers are subsequently able to reap the benefits of their exceedingly high educational attainments during unexpected episodes of increased ambiguity, when less credentialed individuals try to keep up. Such increased uncertainty and turbulence following labor-displacing technological changes is documented by Gottschalk and Moffitt (2009) and Lalé (2018). During such episodes, the affluence of high-income earners – some of which are "superstars" – affords them to optimally invest in their stock of human capital beyond the level which they would under the absence of ambiguity. Put differently, they would have invested less if they knew the true frequency of obsolescence. Correspondingly, low-skill and low-income earners will have a learning deficit. It is important to note that neither the low-skill nor the high-skill are aware of the fact that they are accumulating deficits and gluts in human capital respectively. The educational overinvestment by the high-skill leads to a learning glut which pays large dividends during unexpected transitions that exhibit increased ambiguity, all the while the rest of the population attempts to catch up in educational attainment.
These dynamics are consistent with the empirical findings of Edin et al. (2019). They find that workers at the bottom of their occupations initial earnings distributions suffered considerably larger earnings losses following adoption of labor-displacing technology. They account for these earnings losses by reduced time spent in employment, and increased time in unemployment and retraining. Similarly, in the framework of this current study, given diminishing aversion to ambiguity, less affluent workers cannot afford foregone income preceding obsolescence, and hence fail to prepare for their skills potentially becoming obsolete.

I use specification (15) in Table 6 as the benchmark. Figures 10 and 11 show the composi-

[17] A consequence of keeping beliefs neutral is that positive ambiguity shifts are accompanied by slight optimism, and negative ones by pessimism.

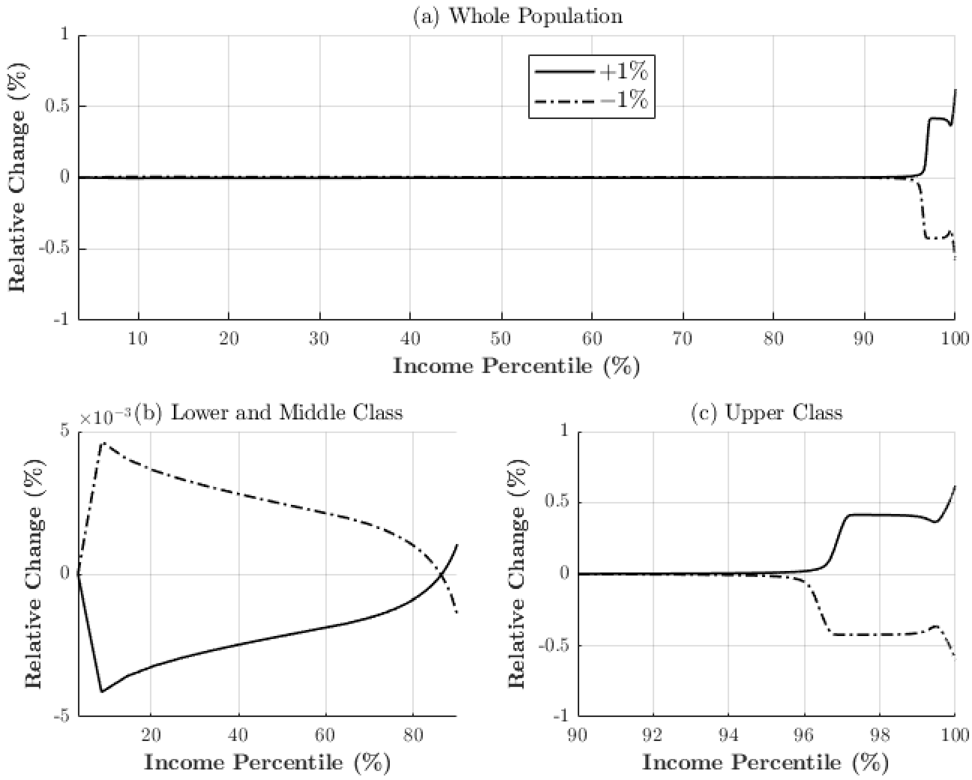

Figure 10: **Equilibrium Changes in Labor Composition following Shifts in Ambiguity.** Specification (15) in Table 6 is used as the benchmark. The figures show the relative change of labor supply of the individuals at each income percentile relative to their own previous supply. In other words, the vertical axis shows relative change in labor supply, while the horizontal axis represents the income percentile.

tional responses of relative ambiguity shifts ($\frac{\Delta\xi}{\xi}$) with magnitudes plus and minus one. The figures show the change of labor and learning supply of the individuals at each income percentile relative to their own previous supply. Subfigures (a) show the impact for the whole population, (b) for the Lower and middle classes, and (c) for the upper class. I will discuss mainly the dynamics of a positive ambiguity shift. The dynamics of a negative one is merely the reverse.

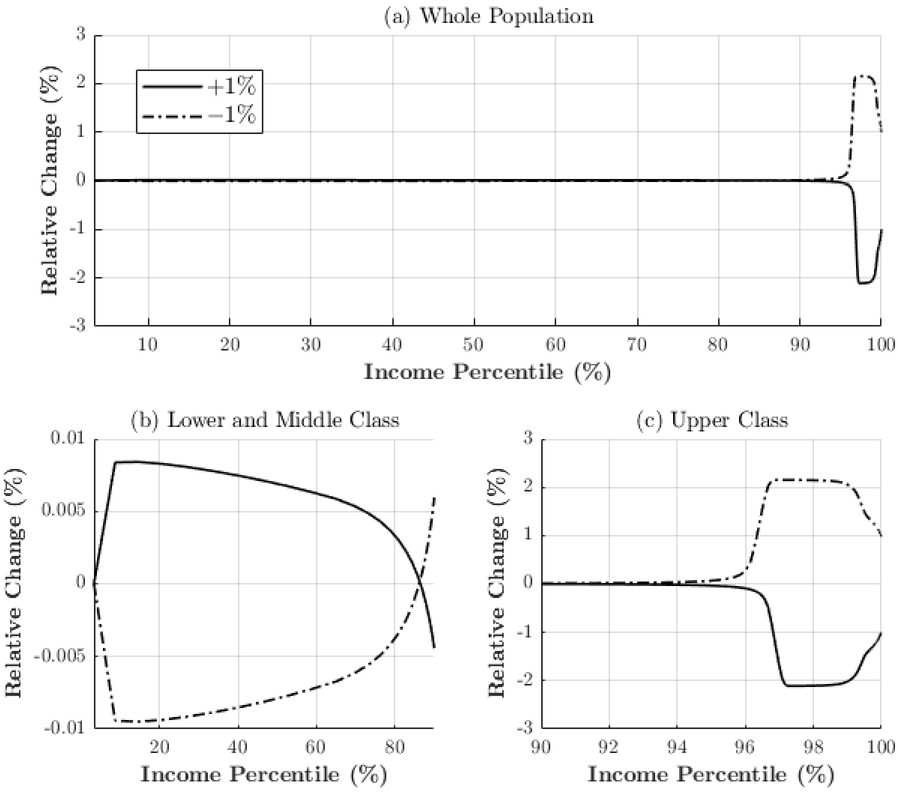

Figure 11: **Equilibrium Changes in Learning Composition due to Shifts in Ambiguity.** Specification (15) in Table 6 is used as the benchmark. The figures show the change of attention to learning of the individuals at each income percentile relative to their own previous supply. In other words, the vertical axis shows relative change in attention to learning, while the horizontal axis represents the income percentile.

The illustrations match the preceding dynamic described. Upon a positive shift in ambiguity, top earners increase their labor supply at the expense of human capital accumulation through learning. These include the whole of the upper class and parts of the upper middle class beyond the 85^{th} percentile (Subfigures (c) in 10 and 11). Even among this group the shift is unequal, with more affluent top earners increasing their labor supply much more than less affluent top earners. At the same time, the rest of the population is trying to catch up by decreasing their labor supply and increasing their educational attainments (Subfigures (b) in 10 and 11). However, as they still cannot afford to forgo that much labor income compared to top earners (Subfigure (a) in 10), their increase in learning is negligible relative to changes for top earners (Subfigure (a) in 11).

An important issue to note is the fact that optimal attention to learning will retain the pattern in Figure 7. In other words, while top earners will decrease their learning, they will still supply the highest attention to learning among the population beyond the threshold $\underline{h}$. As such they continue to accumulate a learning glut, though at a diminished rate. Given the dynamic complementarity in human capital accumulation, lower and middle classes are thus unable to catch up. Their stature in the distribution is thus cemented further, leading to increased skill premium convexity as documented in Figure 12.

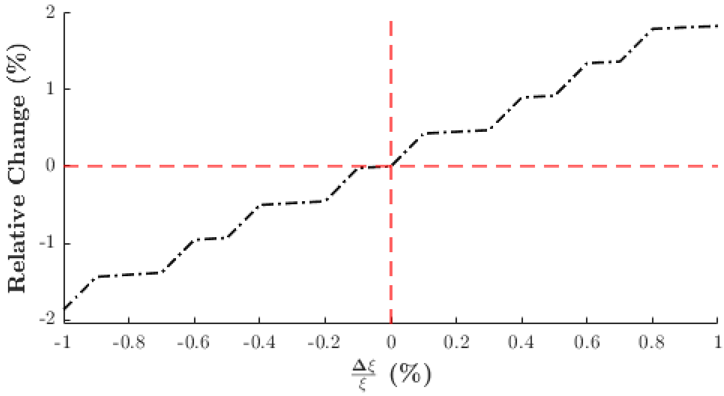

Figure 12: **Equilibrium Changes in Overall Skill Premium Convexity following Shifts in Ambiguity.**
Specification (15) in Table 6 is used as the benchmark. The horizontal axis depcits relative change in ambiguity $\frac{\Delta\xi}{\xi}$ compared to the benchamark. The vertical axis shows the relative change in skill premium convexity as consequence of shifts in ambiguity.

Correspondingly the top earners keep supplying the lowest attention to labor. However, they increase it the most relative to their own previous supply. Hence, their labor income increases significantly. At the same time the rest of the population is engaged in a futile attempt to catch up by learning a bit more. So – in line with Proposition 3.4 – the income share at the top increases, while the middle and lower classes, including the poor lose, . This is depicted in subfigures (a) to (d) of Figure 13. Hence, overall inequality, measured by the Gini coefficient increases as in Figure 13(e). Finally, due to the increase in labor supply at the top, production Y increases, as this increase is larger than at the bottom as seen in Figure 13(f). This is consistent with growth that accompanies periods of increased ambiguity such as adoption of labor-displacing technology or trade deals.

5 Discussion

Some of the findings deserve further discussion. The major contribution of the paper is providing a general account for skill premia's convexity in education. In doing so, it provides microfoundations for increasing and convex structures in premia observed in the data (also cf. e.g. Lemieux, 2006, 2008; Acemoglu and Autor, 2011; Autor, 2014) which have been directly assumed in previous modeling frameworks (Acemoglu and Restrepo, 2018). Under diminishing aversion to ambiguity, affluent workers – colloquially referred to as "superstars" – are shown to invest more in learning as their concern with forgone income due to skill obsolescence is progressively lower than their less credentialed counterparts. This phenomenon is coined as type dependence by Gabaix et al. (2016). This concordance in affluence and learning investments also better captures the top shares of income that are typically underestimated relative to data.

In order to match the empirical pattern of skill premia's convexity in education I applied two

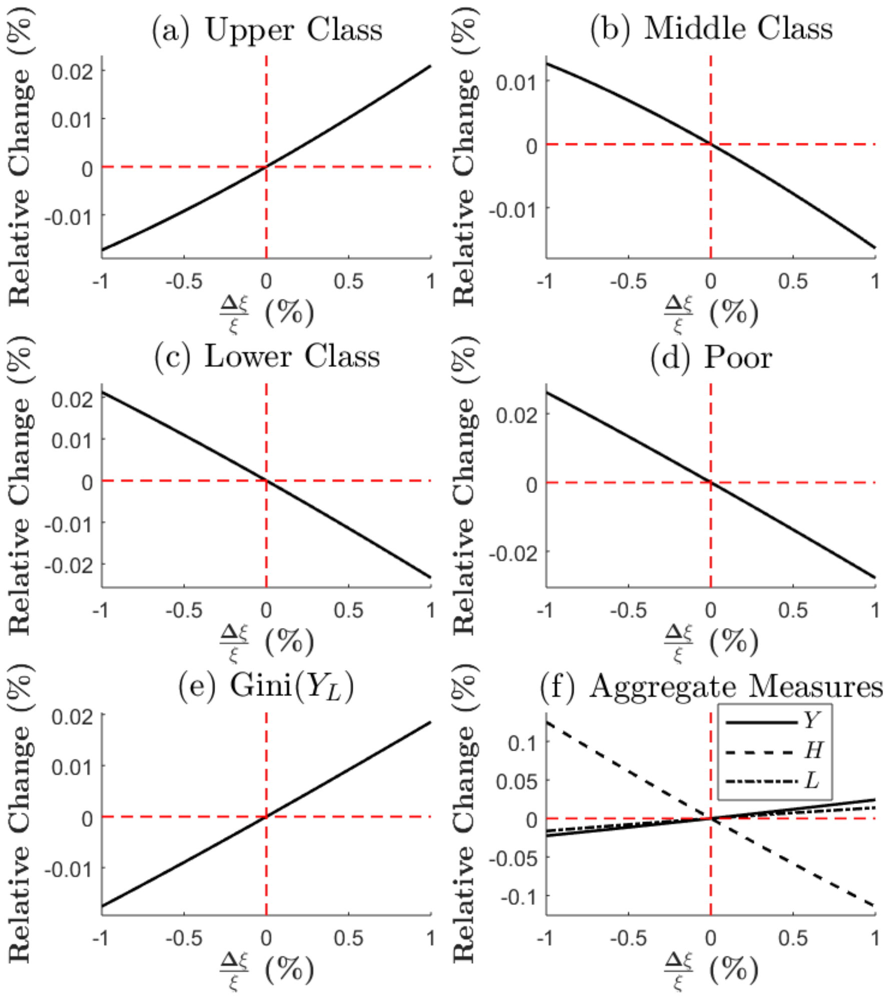

Figure 13: **Equilibrium Changes in Inequality and Aggregate Measures Following Shifts in Ambiguity.**
Specification (15) in Table 6 is used as the benchmark. The horizontal axis in all figures depcits relative change in ambiguity $\frac{\Delta\xi}{\xi}$ compared to the benchamark. The vertical axes in all subfigures shows the relative change in of the mentioned variables in the title as consequence of shifts in ambiguity. In other words, the vertical axis in (a) depicts the relative change in upper class share (top 10 %) while (b), (c) and (d) that of the middle class (median to 90^{th} income 50 percentile), lower class (bottom 50%) and the poor (bottom 10%). The vertical axis in (e) similarly depicts relative change in the Gini coefficient. Finally, the vertical axis (f) shows the relative change in aggregate values of production Y, human capital H and labor supply L.

measures: heightened ambiguity aversion and lower share of firm revenue in wage-setting. The first measure was reasonable since convexity levels were exaggerated when experimentally documented levels of ambiguity aversion were matched. A reasonable conclusion is that aversion levels in the calibration that try to match experimental results of Ahn et al. (2014) are too low. Nevertheless, no burden of error is to be put upon said – and such – studies. More reasonable deduction is there being a discrepancy between measurements in ambiguity aversion at the micro and macro levels akin to the documented discrepancy in measures of risk aversion between the fields (cf. e.g. Chetty, 2006). Reconciling this discrepancy can be done by taking diminishing aversion to ambiguity seriously. In other words, since stakes on a lifetime scale dwarf those of small portfolio selection exercises or betting on Ellsberg urn lotteries, diminishing aversion to ambiguity would entail lower levels in the latter less pressing situations. The second measure also has empirical grounding. In reality, the portion of firm revenue that serves as a basis for wage-setting depends partly on the unions' bargaining power and partly on interfirm competition (cf. e.g. Cahuc, Postel-Vinay and Robin, 2006). The less bargaining power the unions have and the less interfirm competition there exists, the more surplus firms can extract from the workers, and consequently $\vartheta \downarrow 1$, and vice versa. Lowering ϑ scaled down the skill premium curve, in accordance with scale dependence documented by Gabaix et al. (2016).

Once the pattern of skill premia's convexity in education was matched, I could qualitatively show that positive ambiguity shifts are accompanied by increased GDP, inequality and convexity in the skill premium curve. Recent decades have indeed seen such transitional periods which are compatible with increased ambiguity. Recall the intuitive dynamic laid out earlier. Affluent workers tend to invest more in learning as their concern with forgone income due to skill obsolescence is progressively lower than their less credentialed counterparts. The affluent workers are subsequently able to reap the benefits of their exceedingly high educational attainments during transitional episodes of increased ambiguity, when less credentialed individuals try to keep up. Indeed, the affluence of high-income earners – some of which are "superstars" – affords them to optimally invest in their stock of human capital beyond the level which they would under the absence of ambiguity. Put differently, they would have invested less if they knew the true frequency of obsolescence. Correspondingly, low-skill and low-income earners will have a learning deficit. It is important to note that neither the low-skill nor the high-skill are aware of the fact that they are accumulating deficits and gluts in human capital respectively. The educational overinvestment by the high-skill leads to a learning glut which pays large dividends during unexpected transitions that exhibit increased ambiguity, all the while the rest of the population attempts to catch up in educational attainment.

These observations have serious policy implications. Indeed, they suggest that under diminishing aversion to ambiguity, it is optimal for individuals with lower levels of human capital to invest less than their already more educated counterparts, which extenuates the inequality among them. It potentially sheds light on why such differences persist. A typical countervailing measure touted would be subsidized and better quality education. However, such policies will benefit the well-endowed even more, and extenuate their high-growth status. Hence, if the goal of a policy maker is reduced inequality, the analysis suggests these measures either must be directed to the less-educated – excluding the upper class – or be accompanied by some progressive taxation scheme as the ones espoused by Piketty and Saez (2007), Guvenen, Kuruscu and Ozkan (2014) and Ali Akbari and Fischer (2020). Moreover, such measures should be augmented with insights from frameworks that consider a multitude of skill types, such as task-based frameworks in Autor, Levy and Murnane (2003) and Acemoglu and Restrepo (2018).

6 Concluding Remarks

This study provides a theoretical framework linking uncertainty and diminishing aversion towards it to increasing and convex pattern of skill premia as a function of education. The framework uses smooth ambiguity framework based on Klibanoff, Marinacci and Mukerji (2005) – and its continuous-time extension in Skiadas (2013) – that allows for distinguishing between ambiguity, aversion to it (preferences) and attitudes (beliefs). Moreover, I document this framework's strong ability to match distributional moments in the data through calibration exercises. Affluent workers are shown to invest more in education as their concern with forgone income is progressively lower than their less credentialed counterparts. In other words, affluence of high-skill agents – "superstars" among them – affords them a level investment in their stock of human capital beyond what is optimal if true obsolescence frequency was known, i.e. in absence of ambiguity. The emergence of agents with high-skill growth is a phenomenon which has been previously coined as type-dependence by Gabaix et al. (2016). This overinvestment leads to a learning glut which subsequently pays large dividends during unexpected episodes that exhibit increased turbulence and ambiguity, all the while the rest of the population who are experiencing a learning deficit attempts to catch up. Given dynamic complementarity in learning and type dependence, it will be very difficult for lower and middle classes to catch up to the upper class, reducing social mobility. Indeed, as empirical findings of Edin et al. (2019) suggest, workers at the bottom of their occupations initial earnings distributions suffer considerably larger earnings losses following adoption of labor-displacing technology. These earnings losses are accounted for by reduced time spent in employment, and increased time in unemployment and retraining. In this study, I show that given diminishing aversion to ambiguity, less affluent workers cannot afford foregone income preceding obsolescence, and hence fail to prepare for their skills potentially becoming obsolete. Thereby, this study focuses on how workers' alter their attainment, and subsequently, supply of skill in response to said demand shifts.

A number of future investigations come to mind. One is putting these insights to test in a general-equilibrium setting that allows for insuring against labor income shocks through precautionary saving. Another interesting question is the role capital and skill complementarity plays in qualifying skill premium. Finally, this author calls for more research into estimating aversion to ambiguity when stakes are comparable to those that impact lifetime earnings. Workers face ambiguity in their learning decisions. Hence, further investigation into the interplay between second-order uncertainty and skill premium will further deepen our understanding of the emerging patterns of inequality that we see in society writ large – patterns that include increased skill premium convexity in education, job market polarization and larger top shares in income.

References

Abdellaoui, Mohammed, Aurélien Baillon, Laetitia Placido and Peter P Wakker. 2011. "The rich domain of uncertainty: Source functions and their experimental implementation." *American Economic Review* 101(2):695–723.

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, Gino Gancia and Fabrizio Zilibotti. 2012. "Competing engines of growth: Innovation and standardization." *Journal of Economic Theory* 147(2):570–601.

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

Acemoglu, Daron and Pascual Restrepo. 2020. "Robots and jobs: Evidence from US labor markets." *Journal of Political Economy* 128(6):2188–2244.

Ahn, David, Syngjoo Choi, Douglas Gale and Shachar Kariv. 2014. "Estimating ambiguity aversion in a portfolio choice experiment." *Quantitative Economics* 5(2):195–223.

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

Altonji, Joseph G, Erica Blom and Costas Meghir. 2012. "Heterogeneity in human capital investments: High school curriculum, college major, and careers." *Annual Review of Economics* 4(1):185–223.

Aubert, Patrick, Eve Caroli and Muriel Roger. 2006. "New technologies, organisation and age: firm-level evidence." *The Economic Journal* 116(509):F73–F93.

Autor, David and Anna Salomons. 2018. "Is automation labor-displacing? Productivity growth, employment, and the labor share." No. 24871.

Autor, David H. 2014. "Skills, education, and the rise of earnings inequality among the other 99 percent." *Science* 344(6186):843–851.

Autor, David H, David Dorn and Gordon H Hanson. 2016. "The China shock: Learning from labor-market adjustment to large changes in trade." *Annual Review of Economics* 8:205–240.

Autor, David H, 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.

Axtell, Robert L. 2001. "Zipf distribution of US firm sizes." *Science* 293(5536):1818–1820.

Baillon, Aurélien and Han Bleichrodt. 2015. "Testing ambiguity models through the measurement of probabilities for gains and losses." *American Economic Journal: Microeconomics* 7(2):77–100.

Baillon, Aurélien and Lætitia Placido. 2019. "Testing constant absolute and relative ambiguity aversion." *Journal of Economic Theory* 181:309–332.

Barrero, Jose Maria, Nicholas Bloom and Steven J Davis. 2020. "Covid-19 is also a reallocation shock." *NBER Working Paper Series* No. 27137.

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.

Bertola, Giuseppe and Andrea Ichino. 1995. "Wage inequality and unemployment: United States vs. Europe." *NBER Macroeconomics Annual* 10:13–54.

Bhagwati, Jagdish, Arvind Panagariya and Thirukodikaval N Srinivasan. 2004. "The muddles over outsourcing." *Journal of Economic Perspectives* 18(4):93–114.

Cahuc, Pierre, Fabien Postel-Vinay and Jean-Marc Robin. 2006. "Wage bargaining with on-the-job search: Theory and evidence." *Econometrica* 74(2):323–364.

Carroll, Christopher D and Andrew A Samwick. 1998. "How important is precautionary saving?" *Review of Economics and Statistics* 80(3):410–419.

Chen, Peter, Loukas Karabarbounis and Brent Neiman. 2017. "The global rise of corporate saving." *Journal of Monetary Economics* 89:1–19.

Chetty, Raj. 2006. "A new method of estimating risk aversion." *American Economic Review* 96(5):1821–1834.

Collard, Fabrice, Sujoy Mukerji, Kevin Sheppard and Jean-Marc Tallon. 2018. "Ambiguity and the historical equity premium." *Quantitative Economics* 9(2):945–993.

Cubitt, Robin, Gijs Van De Kuilen and Sujoy Mukerji. 2020. "Discriminating between models of ambiguity attitude: A qualitative test." *Journal of the European Economic Association* 18(2):708–749.

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

Davis, Steven J, R Jason Faberman, John Haltiwanger, Ron Jarmin and Javier Miranda. 2010. "Business volatility, job destruction, and unemployment." *American Economic Journal: Macroeconomics* 2(2):259–87.

De Grip, Andries and Jasper Van Loo. 2002. *The Economics of Skill Obsolescence: A Review.* Elsevier Science.

Den Haan, Wouter J, Garey Ramey and Joel Watson. 2000. "Job destruction and propagation of shocks." *American Economic Review* 90(3):482–498.

Ebenstein, Avraham, Ann Harrison, Margaret McMillan and Shannon Phillips. 2014. "Estimating the impact of trade and offshoring on American workers using the current population surveys." *The Review of Economics and Statistic* 96(4):581–595.

Edin, Per-Anders, Tiernan Evans, Georg Graetz, Sofia Hernnäs and Guy Michaels. 2019. "Individual consequences of occupational decline." *CEPR Discussion Paper* No. DP13808.

Ellsberg, Daniel. 1961. "Risk, ambiguity, and the Savage axioms." *The quarterly journal of economics* pp. 643–669.

Epstein, Larry G. 1999. “A definition of uncertainty aversion.” *The Review of Economic Studies* 66(3):579–608.

Gabaix, Xavier. 1999. “Zipf’s law for cities: An explanation.” *The Quarterly Journal of Economics* 114(3):739–767.

Gabaix, Xavier. 2009. “Power laws in economics and finance.” *Annual Review of Economics* 1:255–294.

Gabaix, Xavier. 2016. “Power laws in economics: An introduction.” *Journal of Economic Perspectives* 30(1):185–206.

Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions and Benjamin Moll. 2016. “The dynamics of inequality.” *Econometrica* 84(6):2071–2111.

Ghirardato, Paolo, Fabio Maccheroni, Massimo Marinacci et al. 2004. “Differentiating ambiguity and ambiguity attitude.” *Journal of Economic Theory* 118(2):133–173.

Gilboa, Itzhak. 1987. “Expected utility with purely subjective non-additive probabilities.” *Journal of Mathematical Economics* 16(1):65–88.

Gilboa, Itzhak and David Schmeidler. 1989. “Maxmin expected utility with a non-unique prior.” *Journal of Mathematical Economics* 18:141–153.

Giustinelli, Pamela and Nicola Pavoni. 2017. “The evolution of awareness and belief ambiguity in the process of high school track choice.” *Review of Economic Dynamics* 25:93–120.

Gollier, Christian. 2011. “Portfolio choices and asset prices: The comparative statics of ambiguity aversion.” *The Review of Economic Studies* 78(4):1329–1344.

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.

Gottschalk, Peter and Robert Moffitt. 1994. “The growth of earnings instability in the US labor market.” *Brookings Papers on Economic Activity* 1994(2):217–272.

Gottschalk, Peter and Robert Moffitt. 2009. “The rising instability of US earnings.” *Journal of Economic Perspectives* 23(4):3–24.

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

Gul, Faruk and Wolfgang Pesendorfer. 2014. “Expected uncertain utility theory.” *Econometrica* 82(1):1–39.

Gul, Faruk and Wolfgang Pesendorfer. 2020. “Calibrated Uncertainty.” *Journal of Economic Theory* 188(105016).

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.

Halevy, Yoram. 2007. "Ellsberg revisited: An experimental study." *Econometrica* 75(2):503–536.

Hayashi, Takashi and Jianjun Miao. 2011. "Intertemporal substitution and recursive smooth ambiguity preferences." *Theoretical Economics* 6(3):423–472.

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.

Jensen, Robert. 2010. "The (perceived) returns to education and the demand for schooling." *The Quarterly Journal of Economics* 125(2):515–548.

Ju, Nengjiu and Jianjun Miao. 2012. "Ambiguity, learning, and asset returns." *Econometrica* 80(2):559–591.

Kahn, Matthew E, Kamiar Mohaddes, Ryan NC Ng, M Hashem Pesaran, Mehdi Raissi and Jui-Chung Yang. 2019. "Long-term macroeconomic effects of climate change: A cross-country analysis." *NBER Working Paper Series* No. w26167.

Klibanoff, Peter, Massimo Marinacci and Sujoy Mukerji. 2005. "A smooth model of decision making under ambiguity." *Econometrica* 73(6):1849–1892.

Klibanoff, Peter, Massimo Marinacci and Sujoy Mukerji. 2009. "Recursive smooth ambiguity preferences." *Journal of Economic Theory* 144(3):930–976.

Knight, Frank. 1921. "Risk, ambiguity, and profit." *Boston, MA: Houghton Mifflin* .

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

Krueger, Dirk, Kurt Mitman and Fabrizio Perri. 2016. Macroeconomics and household heterogeneity. In *Handbook of Macroeconomics*. Vol. 2 Elsevier pp. 843–921.

Kuhn, Moritz and José-Víctor Ríos-Rull. 2016. "2013 Update on the US earnings, income, and wealth distributional facts: A View from Macroeconomics." *Federal Reserve Bank of Minneapolis Quarterly Review* 37(1):2–73.

Lalé, Etienne. 2018. "Turbulence and the employment experience of older workers." *Quantitative Economics* 9(2):735–784.

Lemieux, Thomas. 2006. "Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill?" *American Economic Review* 96(3):461–498.

Lemieux, Thomas. 2008. "The changing nature of wage inequality." *Journal of population Economics* 21(1):21–48.

Ljungqvist, Lars and Thomas J Sargent. 1998. "The European unemployment dilemma." *Journal of Political Economy* 106(3):514–550.

Ljungqvist, Lars and Thomas J Sargent. 2008. "Two questions about European unemployment." *Econometrica* 76(1):1–29.

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

Maccheroni, Fabio, Massimo Marinacci and Aldo Rustichini. 2006. "Ambiguity aversion, robustness, and the variational representation of preferences." *Econometrica* 74(6):1447–1498.

McDonald, Ian M and Robert M Solow. 1981. "Wage bargaining and employment." *American Economic Review* 71(5):896–908.

Miao, Jianjun, Bin Wei and Hao Zhou. 2019. "Ambiguity aversion and the variance premium." *Quarterly Journal of Finance* 9(02):1950003.

Michelacci, Claudio and David Lopez-Salido. 2007. "Technology Shocks and Job Flows." *Review of Economic Studies* 74(4):1195–1227.

Millner, Antony, Simon Dietz and Geoffrey Heal. 2013. "Scientific ambiguity and climate policy." *Environmental and Resource Economics* 55(1):21–46.

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

Mortensen, Dale T and Christopher A Pissarides. 1994. "Job creation and job destruction in the theory of unemployment." *The Review of Economic Studies* 61(3):397–415.

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

Peysakhovich, Alexander and Jeffrey Naecker. 2017. "Using methods from machine learning to evaluate behavioral models of choice under risk and ambiguity." *Journal of Economic Behavior & Organization* 133:373–384.

Piketty, Thomas. 2014. *Capital in the twenty-first century.* Belknap Press of Harvard University Press.

Piketty, Thomas and Emmanuel Saez. 2007. "How progressive is the US federal tax system? A historical and international perspective." *Journal of Economic Perspectives* 21(1):3–24.

Piketty, Thomas and Gabriel Zucman. 2014. "Capital is back: Wealth-income ratios in rich countries 1700–2010." *The Quarterly Journal of Economics* 129(3):1255–1310.

Prelec, Drazen. 1998. "The probability weighting function." *Econometrica* pp. 497–527.

Saez, Emmanuel and Gabriel Zucman. 2016. "Wealth inequality in the United States since 1913: Evidence from capitalized income tax data." *The Quarterly Journal of Economics* 131(2):519–578.

Schmeidler, David. 1989. "Subjective probability and expected utility without additivity." *Econometrica* 57(3):571–587.

Segal, Uzi. 1987. "The Ellsberg paradox and risk aversion: An anticipated utility approach." *International Economic Review* 28(1):175–202.

Shimer, Robert. 2005. "The cyclical behavior of equilibrium unemployment and vacancies." *American Economic Review* 95(1):25–49.

Skiadas, Costis. 2013. "Smooth ambiguity aversion toward small risks and continuous-time recursive utility." *Journal of Political Economy* 121(4):775–792.

Toda, Alexis Akira. 2012. "The double power law in income distribution: Explanations and evidence." *Journal of Economic Behavior & Organization* 84(1):364–381.

Toda, Alexis Akira and Kieran Walsh. 2015. "The double power law in consumption and implications for testing Euler equations." *Journal of Political Economy* 123(5):1177–1200.

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.

Zucman, Gabriel. 2019. "Global wealth inequality." *Annual Review of Economics* 11:109–138.

Appendix

Proofs

Proof of Theorem 3.1. The agent has the following HJB-equation:

$$\rho V(h) = \max_a \left\{ \ln\left(w(1-a)h\right) + V'(h)g(a)h + \frac{1}{2}\sigma^2 h^2 V''(h) + \left(\lambda + \lambda^u u(h)\right)\left(V(h^o) - V(h)\right) \right\}$$

We have dropped the indexes just for notational compactness where $V(h)$ is the optimal value function. The first-order condition of the right-hand side is given by:

$$\frac{1}{1-a} = V'(h)g'(a)h. \tag{41}$$

We guess the solution as $V(h) = A + B\ln(h)$ where $A, B \in \mathbb{R}$. Observe that $V(h^o) = A + B\ln(1) = A$. Inserting this guess into the FOC and then the HJB equation we get:

$$\rho A + \rho B \ln h = \ln(1-a) + \ln w + B\left(g(a) - \lambda^u \cdot 1_{\{h>h^o\}} - \frac{1}{2}\sigma^2\right) + (1 - \lambda B)\ln h$$

Hence,

$$B = \frac{1}{\rho + \lambda} \tag{42}$$

$$A = \frac{1}{\rho}\left(\ln(1-a) + \ln w + \frac{1}{\rho+\lambda}(g(a) - \lambda^u \cdot 1_{\{h>h^o\}} - \frac{1}{2}\sigma^2)\right) \tag{43}$$

Then the FOC in (41) is given by:

$$\frac{1}{1-a^*} = \frac{g'(a^*)}{\rho+\lambda}. \tag{44}$$

which directly shows independence of the optimal choice of attention a^* from level of human capital h and wage rate $\bar{w}$, i.e. $\frac{d}{dh}a^* = 0$ and $\frac{d}{dw}a^* = 0$, where a^* is the control that maximizes the objective function. Differentiating both sides of (44) with respect to the frequency of obsolescence λ results in the following expression:

$$\frac{d}{d\lambda}a^* = \frac{-g'(a^*)}{\left(\frac{\rho+\lambda}{1-a^*}\right)^2 - g''(a^*)} < 0 \tag{45}$$

which is negative since $g' > 0$ and $g'' < 0$. Similar derivations show that $\frac{d}{d\rho}a^* < 0$. □

Proof of Corollary 3.1.2. Observe that the optimal value function is given by $V(h) = A + B\ln(h)$ where B and A as in (42) and (43) respectively. Differentiating $V(h)$ with respect to λ and applying the envelope theorem we get:

$$\frac{\partial}{\partial\lambda}V = -\frac{1}{(\rho+\lambda)^2}\left(\rho\ln h + g(a^*) - \lambda^u \cdot 1_{\{h>h^o\}} - \frac{1}{2}\sigma^2\right) \tag{46}$$

which proves the corollary. □

Lemma 6.1. *Define $a_s^* \equiv a^*(\lambda_s)$ and $a_l^* \equiv a^*(\lambda_l)$. Then*

$$a^{**}(\lambda^*) \in [a_l^*, a_s^*], \qquad \textit{and} \qquad \lambda^* \in [\lambda_s, \lambda_l]. \tag{47}$$

Proof. We first show that $a^{**} \in [a_l^*, a_s^*]$ as proof by contradiction: Assume, first $a^{**} > a_s^*$. However, then both $\phi(V(h;\lambda_s, a^{**}))$ and $\phi(V(h;\lambda_l, a^{**}))$ in the objective function (13) can be increased by lowering the amount of attention to learning a^{**}, which contradicts the fact that a^{**} is optimizing. Similarly, if $a^{**} < a_l^*$, then the same expressions can be increased by elevating the optimal amount of attention to learning a^{**}. It follows then trivially that $\lambda^* \in [\lambda_s, \lambda_l]$, since $\frac{d}{d\lambda}a^* < 0$ by Proposition 3.1. □

Lemma 6.2. *Assume that $h > \tilde{h}(a^{**})$ and that the optimal amount of attention to learning is an interior point, $a^{**} \in (a_l^*, a_s^*)$. Then less strenuous best-case scenario and more strenuous worst-case scenario in frequency of incumbent skill obsolescence leads to more human capital* *investment if and only if the agent has low enough aversion towards ambiguity. More precisely*

$$\text{(a) } \frac{\partial}{\partial \lambda_l} a^{**} < 0 \qquad \text{and} \qquad \text{(b) } \frac{\partial}{\partial \lambda_s} a^{**} > 0 \tag{48}$$

where the comparative statics (a) has only (12) as necessary and sufficient condition, while (b) requires further that

$$A(V_s) < \frac{\frac{\partial^2}{\partial a \partial \lambda_s} V_s}{\frac{\partial}{\partial a} V_s \cdot \frac{\partial}{\partial \lambda_s} V_s}. \tag{49}$$

If $h \leq \tilde{h}(a^{**})$, then (48b) holds unconditionally.

Proof of Lemma 6.2. Since $a^{**} \in (a_l^*, a_s^*)$, the optimal choice of attention to learning satisfies the following FOC:

$$\frac{\partial}{\partial a} \mathbb{E}_\mu \left(\mathcal{L}(\Lambda, a^{**}) \right) = 0$$

which is given by differentiating (13) with respect to a^* resulting in the first-order condition (16). Observe that since $a^{**} < a_s^*$ by Lemma 6.1, it holds that

$$\frac{\partial}{\partial a} V_s > 0$$

and similarly since $a^{**} > a_l^*$

$$\frac{\partial}{\partial a} V_l < 0.$$

We also know that the agent is ambiguity averse, i.e. $\phi' > 0$ and $\phi'' < 0$. We note that:

$$\frac{\partial^2}{\partial a^2} V = \frac{1}{\rho} \left(\frac{g''(a)}{\rho + \lambda} - \frac{1}{(1-a)^2} \right) < 0$$

regardless of λ. To establish the comparative statics of a^{**}, we differentiate the FOC (16) with respect to λ_l and λ_s.

We subsequently differentiate (16) with respect to λ_l, which yields:

$$\underbrace{\left(p \left[\phi''(V_s) \left(\frac{\partial}{\partial a} V_s \right)^2 + \phi'(V_s) \frac{\partial^2}{\partial a^2} V_s \right] + (1-p) \left[\phi''(V_l) \left(\frac{\partial}{\partial a} V_l \right)^2 + \phi'(V_l) \frac{\partial^2}{\partial a^2} V_l \right] \right)}_{-} \frac{\partial}{\partial \lambda_l} a^{**}$$

$$+ \underbrace{(1-p) \left[\phi''(V_l) \frac{\partial}{\partial a} V_l \frac{\partial}{\partial \lambda_l} V_l + \phi'(V_l) \frac{\partial^2}{\partial \lambda_l \partial a} V_l \right]}_{-} = 0$$

where $\frac{\partial^2}{\partial \lambda_l \partial a} V_l = -\frac{1}{\rho} \frac{g'(a^{**})}{(\rho + \lambda_l)^2} < 0$. Hence, (48b) follows.

Finally we differentiate (16) with respect to λ_s, which yields:

$$\underbrace{\left(p\left[\phi''(V_s)\left(\frac{\partial}{\partial a}V_s\right)^2+\phi'(V_s)\frac{\partial^2}{\partial a^2}V_s\right]+(1-p)\left[\phi''(V_l)\left(\frac{\partial}{\partial a}V_l\right)^2+\phi'(V_l)\frac{\partial^2}{\partial a^2}V_l\right]\right)}_{-}\frac{\partial}{\partial\lambda_s}a^{**}$$

$$+\underbrace{p\left[\phi''(V_s)\frac{\partial}{\partial a}V_s\frac{\partial}{\partial\lambda_s}V_s+\phi'(V_s)\frac{\partial^2}{\partial\lambda_s\partial a}V_s\right]}_{<0\text{ if and only if (49) holds.}}=0$$

where $\frac{\partial^2}{\partial\lambda_s\partial a}V_s=-\frac{1}{\rho}\frac{g'(a^{**})}{(\rho+\lambda_s)^2}<0$. Hence, (48c) follows as well. □

Observe that the left-hand side in (49) is ϕ's *absolute ambiguity aversion*, which is the reason why (49) is interpreted as the agent having low enough ambiguity aversion in Proposition 6.2. If the absolute ambiguity aversion is high enough, then $a^{**}=a_s^*$ with $\frac{\partial}{\partial a}\mathbb{E}_\mu\left(\mathcal{L}(\Lambda,a^{**})\right)<0$, so that a more than marginal increase in the frequency of best-case obsolescence λ_s is needed for the comparative statics in (48b) to apply. Using Lemma (6.2) we can now prove Proposition 3.2.

Remark 6.0.1. *We can readily prove that (48) holds for the end points $a^{**}\in\{a_l^*,a_s^*\}$, if (16) holds, i.e. if they are stationary points to $\mathbb{E}_\mu\left(\mathcal{L}(\Lambda,a^*(\lambda^*))\right)$. If they are not stationary points, then a larger than marginal change in p, λ_l and λ_s is needed to observe the comparative statics indicated by (48) for the end points.*

Proof of Proposition 3.2. To establish the comparative statics of a^{**}, we differentiate the FOC (16) with respect to p and ξ.

First, we differentiate (16) with respect to p, which yields:

$$\underbrace{\phi'(V_s)\frac{\partial}{\partial a}V_s-\phi'(V_l)\frac{\partial}{\partial a}V_l}_{+}+\underbrace{\left(p\left[\phi''(V_s)\left(\frac{\partial}{\partial a}V_s\right)^2\right]+(1-p)\left[\phi''(V_l)\left(\frac{\partial}{\partial a}V_l\right)^2\right]\right)}_{-}\frac{\partial}{\partial p}a^{**}=0$$

Hence, (17a) follows.

In order to prove (17b) note that,

$$\frac{\partial}{\partial\xi}a^{**}=\frac{\partial a^{**}}{\partial\lambda_l}\cdot\frac{\partial\lambda_l}{\partial\xi}+\frac{\partial a^{**}}{\partial\lambda_s}\cdot\frac{\partial\lambda_s}{\partial\xi}=\frac{\partial a^{**}}{\partial\lambda_l}-\frac{\partial a^{**}}{\partial\lambda_s}<0$$

if and only if

$$\frac{\partial a^{**}}{\partial\lambda_l}<\frac{\partial a^{**}}{\partial\lambda_s}. \tag{50}$$

As $\frac{\partial a^{**}}{\partial\lambda_l}<0$, then a sufficient condition for (50) is that $\frac{\partial a^{**}}{\partial\lambda_s}>0$, which by Lemma 6.2 holds when the condition in (18) is met. □

Proof of Proposition 3.3. A proof for the Gini-coefficient is given in Toda (2012), and an alternative derivation in Ali Akbari and Fischer (2020). The proof for expression pertaining to top shares can be found in Appendix C of Ali Akbari and Fischer (2020). □

Proof of Proposition 3.4. We start by deriving the comparative statics of the upper tail $\varsigma_+(h; a^{**}(\lambda^*), \lambda^*)$ with respect to the revealed frequency of obsolescence λ^* showing it to be negative. The rest then follows from Corollary 3.2.1.

Differentiating the upper tail $\varsigma_+(h; \lambda^*)$ with respect to the revealed frequency of obsolescence λ^* we get:

$$-\frac{\partial}{\partial \lambda^*}\varsigma_+(h;\lambda^*) = -\frac{2g'(a^{**})}{\sigma^2}\cdot\frac{\partial a^{**}}{\partial \lambda^*} + \frac{q'(\lambda^*)}{2\sqrt{q(\lambda^*)}}, \tag{51}$$

where

$$q(\lambda^*) = \left(1 - \frac{2g(a^*)}{\sigma^2}\right)^2 - 8\left(1 - \frac{g(a^*) + \lambda^* + \frac{\lambda^u}{\ln h}}{\sigma^2}\right)$$

and hence,

$$q'(\lambda^*) = (2g(a^*) + 1)\cdot\frac{4g'(a^{**})}{\sigma^4}\cdot\frac{\partial a^{**}}{\partial \lambda^*} + \frac{8}{\sigma^2}. \tag{52}$$

Recall that $\frac{\partial a^{**}}{\partial \lambda^*} < 0$ by definition of revealed frequency of obsolescence and that $g' > 0$. So that $-\frac{2g'(a^{**})}{\sigma^2}\cdot\frac{\partial a^{**}}{\partial \lambda^*} > 0$ in (51). So $-\frac{\partial}{\partial \lambda^*}\varsigma_+(h;\lambda^*) > 0$ when $q'(\lambda^*) > 0$, which holds if

$$\sigma^2 > \tilde{\sigma}^2 \equiv \left(g(a^*) + \frac{1}{2}\right)\cdot g'(a^{**})\cdot\left|\frac{\partial a^{**}}{\partial \lambda^*}\right|. \tag{53}$$

Finally, we observe that

$$\frac{\partial}{\partial m}\varsigma_+(h;\lambda^*) = \frac{\partial \varsigma_+(h;\lambda^*)}{\partial \lambda^*}\cdot\frac{\partial \lambda^*}{\partial m} \qquad \text{for} \qquad m \in \{p, \xi\}. \tag{54}$$

Hence, under (53), the comparative statics in (28) follows from Corollary 3.2.1. □