Occupation Mobility, Human Capital and the Aggregate Consequences of Task-Biased Innovations*

Maximiliano Dvorkin† Alexander Monge-Naranjo†§

October 31, 2019

Abstract

We construct a dynamic general equilibrium model with occupation mobility, human capital accumulation and endogenous assignment of workers to tasks to quantitatively assess the aggregate impact of automation and other task-biased technological innovations. We extend recent quantitative general equilibrium Roy models to a setting with dynamic occupational choices and human capital accumulation. We provide a set of conditions for the problem of workers to be written in recursive form and provide a sharp characterization for the optimal mobility of individual workers and for the aggregate supply of skills across occupations. We craft our dynamic Roy model in a production setting where multiple tasks within occupations are assigned to workers or machines. We solve for the balanced-growth path and characterize the aggregate transitional dynamics ensuing task-biased technological innovations. In our quantitative analysis of the impact of task-biased innovations in the U.S. since 1980, we find that they account for an increased aggregate output in the order of 75% and for a much higher dispersion in earnings. If the U.S. economy had higher barriers to mobility, it would have experienced less job polarization but substantially higher inequality and lower output as occupation mobility has provided an "escape" for the losers from automation.

Keywords: Dynamic Roy models, automation, human capital, aggregation.
JEL Classification: E24, E25, J23, J24, J62, O33.

*First draft: February 2019. Previously circulated as: Occupational Choice and the Dynamics of Human Capital, Inequality and Growth. We thank our discussants, Natalia Kovrijnykh and Todd Schoellman for detailed comments and suggestions. In addition, we would like to thank Ariel Burstein, Paco Buera, Jonathan Eaton, Gueorgui Kambourov, Loukas Karabarbounis, Philipp Kircher, Rody Manuelli, Guido Menzio, Giuseppe Moscarini, Pascual Restrepo, Paulina Restrepo, Andres Rodriguez-Clare, Esteban Rossi-Hansberg, Kim Ruhl, Felipe Saffie, Juan Sanchez, Edouard Schaal, Yongs Shin, Gustavo Ventura, and seminar and conference participants for helpful comments. Asha Bharadwaj and Qiuhan Sun provided excellent research assistance. All views and opinions expressed here are the authors' and do not necessarily reflect those of the Federal Reserve Bank of St. Louis or the Federal Reserve System. †*Federal Reserve Bank of St. Louis.* §*Washington University in St. Louis.*

1 Introduction

While expanding aggregate production, technological and organizational advances are usually biased against some occupations and, often, against many workers. For instance, in recent years automation and routinization have received enormous attention because of the displacement of workers from traditional jobs, the polarization of earnings, and a widening inequality in overall income distribution.[1] These "task-biased" technological innovations replace workers for machines in the performance of production tasks, and their aggregate impact depends on how easy or difficult is to substitute factors in the production of different tasks (as emphasized by Autor et al. (2003)) and also on how easy or difficult it is for workers to switch tasks and labor markets (as emphasized by Acemoglu and Autor (2011)). Yet, the ultimate impacts on aggregate output, income distribution and welfare are also determined by the dynamic responses in aggregate investments in equipment, the reallocation of existent and new workers and their human capital, and the equilibrium prices of skills across the different labor market occupations.

In this paper, we use a dynamic general equilibrium model with occupation mobility and endogenous assignment of workers to tasks to quantitatively assess the impact of automation and other task-biased technological innovations. We extend recent quantitative general equilibrium Roy models to a setting with dynamic occupational choices and human capital accumulation. Analytically and quantitatively, we show that these are crucial aspects for the determination of a country's earnings distribution and level of aggregate output, as well as the responses to economy-wide technological advances. In our model, workers have long stochastic lifetimes, are forward looking and can switch occupations in any period during their participation in labor markets. Workers' occupation mobility choices maximize their expected lifetime utility, and, along with idiosyncratic occupation-specific productivity shocks and the costs of switching occupations, drive their accumulation of human capital over the life cycle. With respect to the demand for labor, the production of final goods is the result of assigning different types of workers and machines to tasks across multiple occupations. We show that the equilibrium assignment of workers and machines to tasks generates a (nested) CES aggregate production function that exhibits cross-occupation differences in productivity levels and the complementarity or substitutability between machines and workers.

We quantitatively assess the impact of automation and other task-biased technological innovations on the welfare of workers on different occupations and on the overall earnings inequality for the U.S. economy from 1980 to the present. Mapping our model to the moments observed in the 1970s for the U.S. economy, we then account for the subsequent changes observed in the employment shares across occupations and the increase in earnings inequality that arise from task biased technological advances. An important change observed in U.S. labor markets in the past few decades is the polarization of skills in the labor market, i.e., the decline of employment in

[1] See Acemoglu and Restrepo (2019); Autor and Dorn (2013); Autor, Katz, and Kearney (2006); Autor, Levy, and Murnane (2003); Goos and Manning (2007); Goos, Manning, and Salomons (2014).

middle-skill occupations, like manufacturing and production occupations, and the growth of employment in both high and low-skill occupations, such as managers and professional occupations on one end, and assisting or caring for others on the other.

Using our model we show how task-biased technical improvements can jointly explain the increase in the capital share of income, labor market polarization, earnings inequality and occupational mobility in U.S. labor markets. In our quantitative analysis of the impact of task-biased innovations in the U.S. since 1980, we find that they account for an increased aggregate output in the order of 75% and with much higher dispersion in earnings. We also argue that if the U.S. economy had larger barriers to mobility, it would have experienced less job polarization but would have substantially increased the observed levels of inequality as occupation mobility provides a "escape from the losses from automation. Finally, the model can easily explain the fall in the labor income share: Instead of falling by 9%, the differences in the elasticity of substitution between machines and workers would lead to a fall in this share in the order of 14%.

Our dynamic model highlights the long-lasting impact of task-biased technological changes. Indeed, in our dynamic setting, once-and-for-all changes in automation or other technological changes can lead to sustained growth effects in the earnings of workers over their life-cycle and very long transitional dynamics for the aggregate economy. We emphasize that the welfare and inequality implications for technological changes can be vastly richer than those obtained in other settings as they originate not only from changes in skill prices in each period but also from changes in the equilibrium growth rate of earnings. On the one hand, the positive impact on some workers is not only due to higher level of earnings but also from a faster growth. On the other hand, some workers can be worse-off due to lower levels of earnings and the human capital losses associated with changing occupations at a higher rate. These are aspects that are naturally incorporated in our dynamic setting and are absent from static occupation choice models.

As we further discuss in Section 3, it is extremely challenging to solve for a dynamic general equilibrium Roy model with a non-trivial number of occupations and human capital that endogenously evolves with the workers' occupational choices.[2] In this vein, the paper has a number of expansive methodological contributions. First, we extend the static models used in recent general equilibrium quantitative papers to a recursive setting and fully characterize the solution for the problem of a worker with labor market opportunity shocks that every period may affect his comparative advantage across occupations. We use standard dynamic preferences with constant relative risk-aversion (CRRA) preferences, bridging recent quantitative assignment Roy models (discussed below) with the standard dynamic household model for households used in macroeconomics. Interestingly, by combining CRRA preferences and Frechet distributed labor market

[2]First, the problem of the worker depends on a vector that denotes the worker's skills in each occupation. With a medium to large number of occupations, this is a high-dimensional object that evolves over time, rendering the recursive problem of the worker intractable. Second, solving for general equilibrium requires an aggregation of individual labor supplies comprised of heterogeneous skill levels for all occupations and, in a non-stationary environment, requires the characterization of the dynamic evolution of the aggregate labor supplies in all these markets.

opportunities, we show that the resulting distribution for the continuation values in the Bellman equation of workers follows one of the three extreme value distributions, Frechet, Gumbel, or Weibull, depending on whether the coefficient of relative risk aversion is lower than, equal to, or higher than 1, respectively. In all these cases, conditions for existence and uniqueness are provided and the simple recursion formulas that come out from the Bellman equation make for trivial computations. In doing so, the dynamic problem of workers generates occupation mobility probabilities, not just unconditional occupation choices.

Second, we fully characterize the limiting behavior for the employment and human capital of workers as implied by their dynamic occupation choices. From the worker's individual problems, we derive the law of motion for the employment shares of workers across occupations. Associated with any positive vector of skill prices, we show that there exists a unique invariant distribution of workers. For aggregate human capital, we also show that a simple aggregation property holds, which allows us to write down the transition matrix for the vector of aggregate human capital across occupations. We show that the human capital for each of the cohorts inside a country does not settle down to an invariant state. Instead, each cohort's average grows over time. Using the fact that the dominant eigenvalue of the aggregate human capital transition matrix is always unique, positive, and real, we derive simple formulas for the aggregate human capital of the country as a whole. Thus, our dynamic Roy model explicitly uncovers the simple mechanics by which the life-cycle gradient of earnings of workers is determined by occupation mobility and how it affects the human capital of a country and its long run income inequality within and across cohorts of workers.

Third, by embedding our dynamic Roy model in a fairly rich general equilibrium environment, we develop a framework in which different tasks across occupations are endogenously allocated to workers or machines. In the model, there are two forms of physical capital. The first is the traditional one in neoclassical models, and hence complementary to all workers. The second capital is in the form of 'machines' which compete with –and substitute for– workers. In our setting, output is produced by performing a large set of tasks which are assigned to either workers or machines depending on their relative productivity and relative costs. The productivities of workers relative to their market price and how they compare with those of machines determine the set of tasks they perform within each occupation. From the equilibrium tasks-workers assignment we show that a transparent and tractable nested CES aggregate production function emerges. Importantly, the output labor-share of the economy is an endogenous function of the wages of workers in different occupations, the rental rates of capital, and the labor and machine productivity terms. Fourth, we show the existence of a competitive-equilibrium balanced-growth path for the production economy. Fifth, we extend recent dynamic-hat-algebra methods to models with general CRRA preferences and with human capital accumulation. The advantage of this approach is to substantially reduce the set of parameters needed to be calibrated for the quantitative application of the model. Seventh, we discuss multiple relevant extensions of our baseline model.

The end of this section relates our paper to existing literature. Section 3 briefly overviews the empirical evidence on the impact of task-biased technological innovations on the U.S. labor market. In particular, we highlight the asymmetric impact across workers in different occupations. Section 4 studies the individual optimization problem of a worker that chooses occupations to maximize lifetime utility for a given vector of skill prices or unitary wages per occupation. For given wages, we show the existence of an invariant distribution of workers and human capital across occupations. We also introduce the same demographic structure and derive the occupation choices of new workers. In the last part of Section 4, we provide numerical examples that illustrate the key mechanisms in the model and the implications of technological changes that are biased against routine occupations. Section 5 describes the production economy, solves for the optimal assignment of tasks to workers or machines, derive the aggregate production function and the general equilibrium conditions. That section also describes how to extend the dynamic hat-algebra methods of Caliendo, Dvorkin, and Parro (2019) to our environment with general CRRA preferences and endogenous human capital accumulation. Section 6 contains our quantitative assessment of the impact of task-biased technological innovations and their impact on U.S. labor markets. Section 7 concludes. A number of appendices at the end include the proofs (Appendix A), the dynamic hat algebra (Appendix B), formulae for welfare (Appendix C) and the extension of the model to ex-ante heterogeneous workers (Appendix D.)

Related Literature

Our work is related to a growing but already extensive literature in labor economics and macroeconomics that argues that recent changes in technology have had an asymmetric impact on workers, leading to job polarization and increased earnings inequality. In the careful summary of this literature by Acemoglu and Autor (2011), the authors state that "recent technological developments have enabled information and communication technologies to either directly perform or permit the off-shoring of a subset of the core job tasks previously performed by middle-skill workers, thus causing a substantial change in the returns to certain types of skills and a measurable shift in the assignment of skills to tasks." Using the cross-section of U.S. commuting zones, Autor and Dorn (2013) also find strong empirical support for the asymmetric effects of computerization across occupations and skills. Our paper follows on these authors and uses a task-based framework for analyzing the effect of new technologies on the labor market and their impact on the distribution of earnings. In our model, a worker's human capital evolves endogenously as a result of past labor market decisions. We provide analytic and quantitative results concerning the impact of task biased technologies from the 1970s onward on U.S. workers and their occupational decisions, human capital accumulation, and earnings inequality.

Krusell, Ohanian, Rios-Rull, and Violante (2000) study how skill-biased technical change affect the skill premium and earnings inequality. They argue that the sharp reduction in the price of equipment jointly with capital-skill complementarity differences can account for a significant fraction of the increase in inequality between education groups. In their paper, labor markets are segmented by education, and workers cannot switch across those markets, that is, the type of jobs available for one group of workers is not available for other workers. Our model does not have such a segmentation, as workers can choose between different occupations, trading-off between their own comparative advantage and macroeconomic conditions in these markets. Also, in our model workers accumulate human capital, which they can reallocate across the different occupations. While at the cost of human capital depreciation, occupation mobility provides an "escape" from the adverse effects that technological innovations may have for the earnings in some occupations.

Kambourov and Manovskii (2009) argue that wage inequality and occupational mobility are intimately related. They use a general equilibrium model with occupation-specific human capital and compare economies with different levels of occupational mobility. In our model, workers self-select into occupations according to their implied expected lifetime values, and comparative advantage and relative production costs determine skill prices and workers' allocation. Then, we compute the transition between balanced-growth paths (BGP) from an economy initially as of the late 1970s to the current time and analyze the effects of the technological innovation for the behavior of inequality and growth.

Our paper contributes to the growing quantitative literature that has successfully applied static Roy models with Frechet distributed shocks to diverse topics, and bridges this literature with standard dynamic quantitative macro models.[3] Some of the prominent examples of these papers include the following: Lagakos and Waugh (2013) show that the selection of workers can explain why productivity differences across countries are twice as large in agriculture than outside agriculture. Hsieh, Hurst, Jones, and Klenow (2019) show how discrimination frictions in the labor market across workers with different race and gender have amounted to substantial aggregate misallocation and productivity costs for the U.S. economy as a whole. Galle, Rodríguez-Clare, and Yi (2017) finds that international trade of goods with China increases average welfare, but some groups of workers experience welfare losses as high as five times the average gain. Burstein, Morales, and Vogel (2018) find that the combination of computerization and shifts in occupation demand account for roughly 80% of the rise in the skill premium, with computerization alone accounting for roughly 60%. All in all, our simple recursive methods can be applied to extend this type of models to dynamic contexts, explicitly considering the lifetime implications of staying in a low paying occupation or switching to a better labor market at the expense of a temporary mismatch of their human capital. To be sure, it is straightforward to extend our model to capture ex-ante heterogeneity and age-dependent choices, both of which are salient aspects in the literature on human capital accumulation.

Our paper closely relates to Acemoglu and Restrepo (2018) and Acemoglu and Restrepo (2019) who study how machines and industrial robots affect different workers and labor markets. They

[3]Recent quantitative Roy models build on the original analytic insights of Eaton and Kortum (2002).

argue that this type of technological advance may explain part of the decline in the labor share highlighted in Karabarbounis and Neiman (2013). We extend the assignment model of tasks to workers and machines of Acemoglu and Restrepo (2018) to different factors of production, and in particular different types of workers. We show how our production at the level of each task can be aggregated and derive the equilibrium aggregate production function, which we show is a nested-CES with differences in the elasticity of substitution between capital and labor across different occupations. From here, it is straightforward to characterize the expression for the output labor share, which depends on the equilibrium price of capital, and hence, sensitive to technological innovations that displace labor for capital across some of the tasks. Martinez (2019) obtains a similar aggregation result in production under different assumptions for an economy with cross-firm heterogeneity in automation.

Lee and Wolpin (2006) consider a dynamic equilibrium Roy model, with two production sectors and three labor market skills. Aside from dynamic occupation choices, workers also accumulate human capital. They use their model to explain why the relative constancy of relative wages in face of a substantial expansion of the service sector does not imply that occupation mobility costs are negligible. Indeed, in their counterfactual exercises, they find large income gains if these mobility costs were zero. Similar to them, we find large counterfactual differences arising from occupation mobility costs.

We have abstracted from the rich demographic structure in Lee and Wolpin (2006). Instead, we provide a formulation of the dynamic equilibrium Roy model that leads to a much more tractable aggregation, which greatly facilitates using the model for characterizing the aggregate responses to technology changes, particularly with a large number of occupations. Moreover, we discuss below how we can extend our model to allow for some of the aspects in Lee and Wolpin (2006) and much of the literature on human capital accumulation. Similarly, we discuss how we can extend our model to capture the aggregate equilibrium responses to demographic changes and gender differences, as emphasized by Cortes, Nekarda, Jaimovich, and Siu (2016), and workers' skill and occupational mismatches as emphasized by Guvenen, Kuruscu, Tanaka, and Wiczer (2019).

More recently, Adao, Beraja, and Pandalai-Nayar (2018) also develop a Roy model in which, but they restrict workers to choose their skills before entering labor markets. Once they entered labor markets, their predetermined skills lead workers of different cohorts to choose among two occupations in every period. Thus, workers of different cohorts have different skills and thus select into occupations differently despite facing the same skills prices. Then, they use the model to understand how economies adjust to an asymmetric technological advance, the speed of the adjustment and the impact on income inequality. Instead, our paper emphasizes the evolution of a worker's human capital after the worker has entered the labor markets and how future valuations and not only current payouts determine occupation switches.

Our model highlights how the workers' reallocation across occupations impacts the rate of growth of their earnings over their life-cycle. Recently, Lagakos, Moll, Porzio, Qian, and Schoellman (2018) have documented substantial cross-country differences in the experience-wage profiles, which are on average twice as steep in rich countries as in poor countries. Our analytical aggregation results allow us to transparently examine how technological differences that impact the equilibrium price of skills ultimately lead to differences in the life-cycle wage profiles. We find that at the aggregate level, those differences can imply large permanent differences in aggregate human capital of countries, their aggregate income and inequality. Moreover, we show that once-and-for-all task biased technological innovations may have long lasting effects on the growth rate of the economy.

We use our framework to assess the general equilibrium response of task biased technical change. To this end we use a relatively large number of occupations in our analysis, noting that we are limited by adequate data, not by computing costs. Since the equilibrium aggregate production is a nested-CES with differences in the elasticity of substitution between capital and labor across different occupations, then investment specific technological change Greenwood, Hercowitz, and Krusell (1997) will be non-neutral. Indeed, the resulting response in the model mimics a routine-biased technological as proposed by Autor et al. (2003),Goos and Manning (2007) and Goos et al. (2014), among others. We use our model to quantitatively assess the impact on polarization, income inequality and aggregate human capital and output as workers switch from low-growth occupations to faster growth ones.

To perform these experiments, we apply the recent dynamic-hat-algebra methods of Caliendo, Dvorkin, and Parro (2019) which hugely reduce the number of parameters needed to calibrate or estimate the model and perform quantitative counterfactual experiments. We extend the work of these authors to allow for more general preferences (CRRA instead of log-preferences) and human capital accumulation. Moreover, we show existence (and uniqueness in simple cases) of the competitive general-equilibrium balanced-growth path, describing how the response of human capital accumulation drives the transitional and long-run dynamics after an innovation disrupts the initial equilibrium.

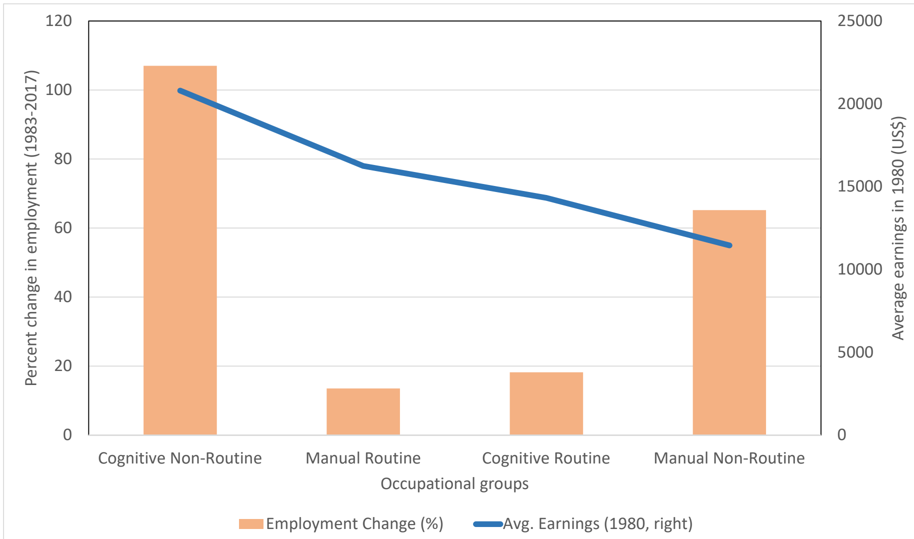
2 Evidence on Task-biased Innovations and Occupations

An extensive literature, partially reviewed above, has emphasized technological advances and the ensuing job polarization and inequality. In this section, we focus on the empirical evidence for the asymmetric effects across different types of occupations, which, following Foote and Ryan (2015), we group into four broad categories: (1) non-routine cognitive, (2) non-routine manual, (3) routine cognitive, and (4) routine manual.[4] Figure 1 shows the cumulative change in employment over the four broad occupational groups between 1980 and 2017. The blue line in the figure (with its scale in the right vertical axis) indicates the average earnings of workers in the four occupations in

[4] In particular, managerial and professional jobs are classified as "non-routine cognitive," service jobs are classified as"non-routine manual," sales and office jobs are classified as "routine cognitive." Finally, construction, repair, production, and transportation jobs are all grouped as "routine manual."

1980. Clearly, during the last four decades, U.S. employment has moved towards the non-routine occupations, which are on the polar extremes of the average earnings distribution, and away from routine occupations, which tend to be in the middle part of the earnings distribution.

Figure 1: Job polarization and earnings 1980-2017



Note: Authors' calculations using data from the Bureau of Labor Statistics and the US decennial census of 1980 (public use microdata).

Specifically, Figure 1 shows that the total employment in non-routine cognitive occupations more than doubled over the last four decades. Not as dramatic but still well above total employment growth, employment in non-routine manual occupations also increased substantially, around 65%. On the contrary, the employment in routine-cognitive and routine-manual occupations, increased much less, both well below 20%. As shares of the total labor, employment in routine occupations has fallen, while employment in non-routine occupations has increased. Under many circumstances, the movement of employment towards the extremes and away from the middle of the pay scale can lead to higher overall earnings inequality.

To explain the shift towards non-routine occupations, recent works have emphasized that technological advances and new forms of capital equipment have displaced workers in routine occupations. This is because many tasks in these occupations tend to be repetitive and easier to codify using a well-defined set of instructions. In turn, these tasks can be easily automatized and performed by machines or software, thus displacing labor from these occupations.

Our empirical analysis adapts the work of Autor and Dorn (2013) and Acemoglu and Restrepo (2019), placing the focus on how technological innovations have impacted different types of occupations. The main idea is to use differences across local labor markets in their exposure to the increase in usage of computers and industrial robots. To the extent that these forms of capital displace workers from some jobs, and, possibly, increase labor demand for others, differences in the characteristics of occupational employment across commuting zones will lead to an above (or below) average exposure in automation and routinization, and an above (or below) average employment changes.

We estimate an empirical equation of the form:

$$\Delta\left(\frac{E_i^{occ\,j}}{Pop_i}\right) = \alpha + \beta\,\Delta\mathrm{PC}_i + \delta\,\Delta\mathrm{APR}_i + \gamma\,X_i + \epsilon_i, \tag{1}$$

where the dependent variable on the left is the change between 1990 to 2007 in the employment of occupational group j, relative to population, for commuting zone i, and the variables $\Delta\mathrm{PC}_i$ and $\Delta\mathrm{APR}_i$ are the changes in the exposure to computers (change in PCs per worker) and robots (adjusted penetration of robots) for commuting zone i, respectively. X_i are controls for commuting zones' characteristics. [5]

We construct the measures for the changes in the exposure to computers and industrial robots at the commuting zone level as follows. For computers, we use the measure of exposure at the commuting zone level from Autor and Dorn (2013).[6] This variable counts the number of personal computers per employee at the firm level constructed by Doms and Lewis (2006), which Autor and Dorn (2013) aggregate to the commuting zone level. For industrial robots, we construct the measures of the change in exposure to robots following closely Acemoglu and Restrepo (2019). We use information from the International Federation of Robotics (IFR), on the stock of robots used in different countries around the world and in different industries by year. In particular, the IFR has data for thirteen industries within manufacturing and for six broad sectors outside of manufacturing. Within manufacturing, we have data on apparel and textiles; automotive; basic metals; clay, glass and minerals; electronics; food and beverages; industrial machinery; metal products; paper and publishing; plastics, chemicals and pharmaceuticals; shipbuilding and aerospace; wood and furniture; and a miscellaneous manufacturing category. Outside of manufacturing, we have data for agriculture, construction, mining, research and education, services, and utilities. Detailed industry-level data for the U.S. is available starting in 2004, which is why our measure for the exposure to robots uses the change in the stock of robots between 2004 and 2007. As in Acemoglu and Restrepo (2019), we re-scale our measure of exposure to robots to match the length of the time period of analysis (1993-2007).[7]

Using information on industry employment for all the United States and for each commuting zone in 1990, together with the stock of robots by industry, we construct the adjusted penetration

[5] Our controls X_i include census division dummies for the commuting zones. All the regressions are weighted by commuting zone population in 1990 and, when necessary, we rescale all variables proportionately to get a 14-year equivalent change (from 1993 to 2007).

[6] This variable was first computed by Doms and Lewis (2006). We take this variable from David Dorn's website: http://www.ddorn.net/data.htm.

[7] We ended our analysis in 2007 to prevent the effects of the Great Recession from confounding our results.

of robots as follows,[8]

$$\Delta \mathrm{APR}_i = \sum_{k=1}^{K} \frac{E_i^k}{E_i} \left(\frac{\Delta M^k}{E^k} - g^k \frac{M^k}{E^k} \right),$$

where E_i^k is the employment in industry k of commuting zone i in 1990, E_i is total employment of the commuting zone in 1990, g^k is the growth rate of output of industry k and M^k is the stock of robots in industry k.

We construct the dependent variable in equation (1) using data from public use microdata sample of the American Census of 1990 and the American Community Survey (ACS). Our sample is restricted to employed individuals between the ages of 25 and 65 that work for a wage or salary (not self-employed) in the private sector. In this way, our sample selection criteria follow closely Acemoglu and Restrepo (2019). The data contains information on the occupation of individuals, which we then aggregate into the four main occupational groups. Because of the potential endogeneity of our regressors, we estimate equation (1) via two-stage least-squares, instrumenting the variables $\Delta \mathrm{PC}_i$ and $\Delta \mathrm{APR}_i$ as follows: First, as Autor and Dorn (2013), we construct an instrument for the change in PCs per worker using industry and occupation information twenty years before the start of the period of analysis. Let $E_{ki,1970}$ equal the employment share of industry $k \in 1, 2, ...K$ in commuting zone i in 1970, and let $R_{k,i,1970}$ equal the routine occupation share among workers in industry k in 1970 in all commuting zones except commuting zone i. The product of these two measures provides a predicted value for the routine employment share in each commuting zone, which depends only on the local industry mix in 1970 and the occupational structure of industries nationally in 1970,

$$R\tilde{S}H_i = \sum_{k=1}^{K} E_{k,i,1970} * R_{k,-i,1970}. \tag{2}$$

Second, we follow Acemoglu and Restrepo (2019) and construct an instrument for the APR in the United States using the changes in the stock of industrial robots in several European countries by industry and using a similar formula as before, but replacing M^k by the stock of robots in European countries in industry k.[9]

Table 1 presents our empirical findings. The data clearly support the hypothesis that the introduction of robots and computers has had diametrically different impacts across occupations. In particular, as robots and computers take over tasks previously performed by workers in routine occupations, the demand for those workers is reduced, while the demand for workers in non-cognitive occupations seems to increase.

First, consider non-routine cognitive occupations. The estimates in the upper-left panel of Table 1 indicates that there is statistically significant evidence that a higher exposure to computers

[8] We use public microdata on employment by industry and commuting zone from the American Census of 1990.

[9] The change in the stock of robots is between 1993 and 2007, and the European countries we use are Denmark, Finland and France.

leads to a higher employment rate in those occupations. This suggests that the productivity of managers and professionals may have increased by the enhanced provision of routine tasks from computers. The positive effect is robust to including robots in the regression. Robots by themselves do not seem to have significant effects on cognitive non-routine occupations.

Table 1: The effects of the increased use of robots and computers on different occupations

	Non-Routine Cognitive			Non-Routine Manual		
	(1)	(2)	(3)	(1)	(2)	(3)
Change in PC's	1.163**		1.142*	0.596		0.585
per worker	(2.59)		(2.55)	(1.69)		(1.71)
Adjusted Exposure		-0.087	-0.082		-0.047	-0.044
to Robots		(-1.61)	(-1.55)		(-0.66)	(-0.66)
observations	660	722	660	660	722	660
Adjusted R^2	0.172	0.164	0.175	0.174	0.170	0.176

	Routine Cognitive			Routine Manual		
	(1)	(2)	(3)	(1)	(2)	(3)
Change in PC's	-3.158***		-3.132***	1.471*		1.314*
per worker	(-6.22)		(-6.25)	(2.28)		(2.19)
Adjusted Exposure		0.118	0.105		-0.622***	-0.616***
to Robots		(0.93)	(1.01)		(-6.04)	(-6.45)
observations	660	722	660	660	722	660
Adjusted R^2	0.170	0.095	0.174	0.283	0.361	0.364

Note: regressions use robust standard errors. t-statistics in parenthesis.

Second, consider the impact on the employment in non-routine manual occupations. The results are reported in the upper-right panel. For those occupations, neither computer nor robots have statistically significant impacts. These results suggest that neither of these forms of technological innovations substitute or complement the workers in non-routine manual occupations, i.e., services, or that, at least at these levels of aggregations, the effects cancel out.

Third, consider the employment in routine manual occupations, reported in the lower-right panel. For these occupations, computers and robots appear to have opposite effects. On the one hand, there is significant evidence that exposure to robots reduces the employment rate in those occupations. On the contrary, although only marginally significant, the exposure to computers seems to enhance the employment in these occupations.

Finally, as shown in the lower-left panel, there is strong evidence that computers reduce the employment rate in routine cognitive occupations. The effects are statistically significant, and robust to the including or not robots, which are not significant. Most interestingly, the magnitude of the effect of computer is the highest across the four groups of occupations.

Overall, these simple results indicate that the introduction of new forms of capital –and the technological innovations embedded in them– have strong and asymmetric impacts across the different occupations. The overarching message is that, on the whole, robots and computers are task-biased innovations that displace routine occupations and possibly favor non-routine occupations. The results suggest an intuitive dichotomy, whereby computers disrupt cognitive occupations and while robots disrupt manual occupations, but both disruptions in the same direction, displacing routine workers and favoring non-routine ones.

The aggregate and individual consequences of task-biased technological innovations crucially depend on how mobile are workers across occupations, and what is the impact of occupation transitions on the human capital of workers. These two margins not only determine the extent of job polarization and earnings inequality, but also, at the macro level, the impact on aggregate output, factor output shares and welfare, and at the individual level, occupational choices and earning dynamics. With those issues in mind, we now construct a quantitative model for the reallocation of workers and their human capital across occupations and use it to examine the aggregate consequences of task-biased technological innovations on employment and factor income shares, aggregate income, inequality and welfare.

3 A Canonical Worker's Problem

We consider an infinite horizon maximization problem for a worker with standard preferences. At any time $t = 0, 1, 2, ...$, the utility of the worker is given by

$$U_t = \frac{(c_t)^{1-\gamma}}{1-\gamma} + E\left[\sum_{s=1}^{\infty} \beta^s \frac{(c_{t+s})^{1-\gamma}}{1-\gamma}\right],$$

where $0 < \beta < 1$ is a discount factor (which accounts for a constant survival probability) and $\gamma \geq 0$ is the coefficient of relative risk aversion (CRRA.) For $\gamma = 1$, we interpret the flow utility to be logarithmic, i.e. $\ln c_t$.

3.1 Dynamic Roy Models: Challenges

The worker starts each period $t = 1, 2, ...$ attached to one of $j = 1, ..., J$ occupations, carrying over from the previous period a vector of human capital x of size J which describes the efficiency units of labor of this worker in each of the J occupations. To set up our framework, we first specify the general dynamic problem of the worker given a time-invariant vector of strictly positive (and finite) wages per unit of human capital $w = \left[w^1, w^2, ... w^\ell ..., w^J\right]$. Therefore, the worker's earnings for the period given her current occupation j are $w^j x_t^j$.

In our model, workers are dynamic optimizers, with their human capital returns as their sole source of income in every period. The worker's consumption in each period is simply the current earnings $w^j x_t^j$. For simplicity, we assume that the evolution of human capital depends on worker's occupational choices, the current level of human capital in all occupations and some random idiosyncratic forces.[10] This minimal set of assumptions allows us to write the problem recursively in the following way,

$$V(j, x_1, x_2, \ldots, x_J) = \frac{(w^j\, x_j)^{1-\gamma}}{1-\gamma} + \beta \max_{\ell} \left\{ E\left[V(\ell, x_1', x_2', \ldots, x_J') \,\middle|\, j, x_1, x_2, \ldots, x_J \right] \right\}$$

This dynamic Roy problem is quite general. It only assumes a first order Markovian dependence of human capital. However, solving for a dynamic general equilibrium model with this level of generality and for a non-trivial number of occupations is extremely challenging. On the one hand, the problem of the worker has the whole vector of human capital as a state variable. With a medium to large number of occupations, this is a high-dimensional object, rendering the problem intractable. On the other hand, solving for the general equilibrium requires an aggregation of individual labor supplies for all markets, which, in a non-stationary environment, requires also the characterization of the dynamic evolution of the aggregate labor supplies in all occupations.

3.2 Main assumptions

Additional assumptions are needed for a tractable dynamic problem that can handle a large number of occupations and make the model suitable for aggregation and general equilibrium. Our assumptions closely connect to recent works on static Roy models with Frechet distributed shocks, extending this literature to a dynamic general equilibrium context.[11]

The two key aspects in a worker's vector of skills, $x \in \mathbb{R}^J$, are his *absolute advantage* and the *comparative advantage* across occupations. The first one is given by the *magnitude* of the vector, $\|x\|$, and defines a common factor on how productive the worker is across the different occupations. The second aspect is determined by the *direction* of the vector, $x(j)/x(k)$, $1 \leq j \neq k \leq J$, and generally determines his best occupation choices.

The first key assumption in our analysis is the *homogeneity* in the utility function of workers and in the law of motion of their human capital. In particular, we assume standard constant relative risk aversion (CRRA) preferences and a linear law of motion for the absolute level of skills of workers.[12] Under these assumptions, we can factor out the absolute advantage $\|x_t\|$ of workers

[10] These assumptions are common in the literature, see for example Lee and Wolpin (2006).

[11] The literature on static general equilibrium Roy models with Frechet distributed shocks has recently been applied to a large number of interesting questions. See for example Burstein et al. (2018); Galle et al. (2017); Hsieh et al. (2019); Lagakos and Waugh (2013).

[12] Log-preferences are also allowed as a special case. See Alvarez and Stokey (1998) for a detailed analysis of dynamic programming problems with homogeneous functions.

and their comparative advantage fully determines the optimal occupation choices for workers.

The second key assumption is a simple *Markov* structure for the comparative advantage of workers. In particular, conditional on worker characteristics and, possibly other pre-determined variables, the current occupation and idiosyncratic random labor market opportunities determine the comparative advantage of workers across occupations. In our model, a matrix τ determines the transferability of skills from each occupation to all others. The diagonal entries of such a matrix govern the average growth in skills in each occupation while the off-diagonal terms capture the different depreciation rates associated with occupation mobility. Moreover, a matrix χ determines non-pecuniary costs or benefits of moving across occupations. In addition to these pecuniary and non-pecuniary costs, in each period, workers realize random labor market opportunities across all occupations. Thus, in each period, the worker's human capital for the next period, in each occupation, is governed by the current occupation, the transferability matrix τ, the absolute level of skills $\|x_t\|$ and a random component for each of the occupations. The absolute level of skills $\|x_t\|$ of workers will grow over time, encoding the worker's entire history of labor market choices and opportunities.

Our assumptions lead to a recursive formulation in which workers optimally self-select across occupations, on the basis of the random opportunities, the human capital transferability and non-pecuniary costs of moving, τ and χ, and endogenously determined valuations of occupations. These valuations solve a fixed point problem that incorporates not only the current prices of skills, but also the future occupation alternatives for workers. At the individual level, the model generates a simple probabilistic model, in which job transition probabilities capture persistence in occupation choices and internalize self-selection and the current and future returns of jobs. At the aggregate level, the model is very tractable for general equilibrium, since it leads to very simple aggregation of the workers and human capital across the different occupations.

3.3 Individual problem

The worker starts each period $t = 1, 2, ...$ attached to one of $j = 1, ..., J$ occupations, carrying over from the previous period an *absolute level* of human capital $h > 0$. Available for the next period, the worker realizes a vector $\epsilon_t = \left[\epsilon_t^1, \, \epsilon_t^2, \, ...\epsilon_t^\ell ...,\epsilon_t^J\right] \in \mathbb{R}_+^J$ of labor market opportunities. Each entry in the vector corresponds to the labor market opportunity in the respective occupation. On the basis of these opportunities, the worker chooses to either stay in the current occupation j or to move to an alternative occupation ℓ.

Switching occupations entails costs (or returns) which we specify as follows: A $J \times J$ *human capital transferability* matrix, with strictly positive entries, $\tau_{j\ell}$, determines the fraction of the human capital h that can be transferred from the current occupation j to a new occupation ℓ. On average, there is depreciation if $\tau_{j\ell} < 1$ or positive accumulation if $\tau_{j\ell} > 1$. The diagonal terms, τ_{jj}, may be higher than one, capturing learning-by-doing, i.e. the accumulated experience capital of a worker as he spends more time in an occupation j. These diagonal terms τ_{jj} may vary by occupation j. The off-diagonal terms $\tau_{j\ell}$ may be less than one to capture a potential mismatch between the human capital acquired in one occupation and the productivity of the worker in a different occupation. Still, some of the off-diagonal terms could be greater than one, capturing skill transferability and cross-occupation training or upgrading. In our specification, these occupation-switching costs have a permanent impact on the human capital of the worker for all future periods and for all future occupation choices. In addition, a $J \times J$ *utility cost* matrix, with strictly positive entries, $\chi_{j\ell}$, captures the non-pecuniary costs (or benefits) of switching from current occupation j to occupation ℓ. These costs shape the occupational decision, but have no direct impact on current or future earnings. We assume these non-pecuniary costs are proportional to the expected lifetime utility of the destination occupation at the time of making the decision.

The human capital of the worker evolves according to the labor market opportunities ϵ_t and the occupation choice of the worker. Given a level of human capital, h, a current occupation j, and a vector $\epsilon_t \in \mathbb{R}_+^J$ of labor market opportunities, the vector

$$h_t\, \tau_{j,\cdot} \odot \epsilon_t \in \mathbb{R}_+^J,$$

describes how many efficient units of labor services, or effective human capital, the worker can provide for each of the alternative occupations $\ell = 1, ..., J$. Here the operator $\odot$ denotes an element-by-element multiplication (Hadamard product). After choosing which occupation to take, the scale of the human capital level for the worker for the next period would be

$$h_{t+1} = h_t \tau_{j_t, \ell_{t+1}}\, \epsilon_{\ell,t}, \tag{3}$$

where j_t and ℓ_{t+1} indicate, respectively, the occupations at period t and $t+1$.[13] It is important to highlight that the role of variable h this model. On the one hand, it is an absolute level of general human capital (or an absolute advantage across occupations). On the other, all the past history of the worker in terms of different occupation choices and realization of idiosyncratic shocks ϵ, is summarized in a single value for h. In this way we sacrifice some generality in favor of tractability.

We now set up the problem of the worker recursively, and provide additional structure to derive a sharp characterization of the optimal choices. Denote by $V(j, h, \epsilon)$ the expected life-time discounted utility of the worker. The Bellman Equation (BE) that defines this value function is,

$$V(j, h, \epsilon) = \frac{\left(w^j\, h\, \epsilon^j\right)^{1-\gamma}}{1-\gamma} + \beta \max_{\ell} \left\{ \chi_{j,\ell}\, E_{\epsilon'} V\left[\ell,\, h',\, \epsilon'\right] \right\}, \tag{4}$$

where $E_{\epsilon'}[\cdot]$ is the expectation over the next period's vector of job market opportunities and h' is given by equation (3.)

To characterize this BE, we first show that it can be *factorized*, i.e. its value can be decomposed

[13] We can use a slightly more general function for human capital accumulation, but, as discussed next, we require it to be homogeneous of degree one in h.

into a factor that depends only on the current occupation and labor market realizations, (j, ϵ), and another factor that depends only on the absolute level of human capital, h. This can be done for any generic distribution for the labor market shocks ϵ for which an expectation satisfies a boundedness condition. *In all what follows, we assume that ϵ is distributed independently over time and across workers, and that all the required moments involving ϵ are finite and well defined.*

First, note that if occupation ℓ is chosen, then, the next period human capital is $h' = h\,\tau_{j\ell}\epsilon_\ell$. Then, observe that the period utility function is homogeneous of degree $1-\gamma$ in h. Therefore, under the hypothesis that the value $V(j,h,\epsilon)$ is homogeneous of degree $1-\gamma$ in h, for any pair (j,ϵ), it can be factorized into a real value $v(j,\epsilon)$ and a human capital factor $h^{1-\gamma}$, i.e., $V(j,h,\epsilon) = v(j,\epsilon)\,h^{1-\gamma}$.[14] Under this hypothesis, the Bellman Equation (4) becomes

$$v(j,\epsilon)\,h^{1-\gamma} = \left(\frac{\left(w^j\,\epsilon^j\right)^{1-\gamma}}{1-\gamma} + \beta\,\max_\ell \left\{\chi_{j,\ell}\,E_{\epsilon'}\left[v(\ell,\epsilon')\right]\left(\tau_{j,\ell}\epsilon^\ell\right)^{1-\gamma}\right\}\right) h^{1-\gamma}.$$

Simplifying out the term $h^{1-\gamma}$ we end up with

$$v(j,\epsilon) = \frac{\left(w^j\,\epsilon^j\right)^{1-\gamma}}{1-\gamma} + \beta\max_\ell\left\{\chi_{j,\ell}\,\left(\tau_{j,\ell}\epsilon^\ell\right)^{1-\gamma} E_{\epsilon'}\left[v(\ell,\epsilon')\right]\right\}, \tag{5}$$

which verifies the factorization hypothesis. Therefore, the characterization of $V(j,h,\epsilon)$ boils down to the characterization of $v(j,\epsilon)$, a random variable that depends on each realization ϵ. For all occupations $j = 1, \ldots J$, denote by v^j the conditional expectation of this random variable, i.e.,

$$v^j \equiv E_\epsilon\left[v(j,\epsilon)\right].$$

Using this definition, and taking the expectation $E_\epsilon[\cdot]$ in both the right- and left-hand sides of (5), the equation reduces to a recursion on v^j:

$$v^j = \begin{cases} \Upsilon^j \frac{\left(w^j\right)^{1-\gamma}}{1-\gamma} + \beta E_\epsilon \max_\ell \left[\left\{\chi_{j,\ell}\,\left[\tau_{j,\ell}\,\epsilon^\ell\right]^{1-\gamma} v^\ell\right\}\right], & \text{for } \gamma \neq 1, \\ \Upsilon^j + \ln w^j + \beta E_\epsilon\left[\max_\ell\left\{v^\ell + \frac{\ln\left(\tau_{j,\ell}\epsilon^\ell\right)}{1-\beta}\right\}\right], & \text{for } \gamma = 1, \end{cases} \tag{6}$$

where $\Upsilon^j = E\left[(\epsilon^j)^{1-\gamma}\right]$ for $\gamma \neq 1$, and $\Upsilon^j = E\left[\log(\epsilon^j)\right]$ if $\gamma = 1$, are scalars that are specific to occupation j.

For all $\gamma \geq 0$, the following lemma establishes simple conditions on the stochastic behavior of the labor market opportunities of workers, that guarantee the existence and uniqueness of values $v \in \mathbb{R}^J$ that solve (6). All along, we assume that $\tau_{j,\ell} > 0$ and $\chi_{j,\ell} > 0$ for all j, ℓ and that the support of ϵ^ℓ is $[0,\infty)$ for all ℓ.

[14] For the logarithmic case, $\gamma = 1$, $V(j,h,\epsilon) = u_j + \beta\left[\max_\ell\left\{v_\ell + \frac{[\ln(h)+\ln(\tau_{j,\ell}\epsilon_\ell)]}{1-\beta}\right\}\right]$, as shown in the Appendix.

Depending on the value of γ, and for each $j = 1, ..., J$, we define the terms, Φ_j as follows:

$$\Phi_j \equiv \begin{cases} E_\epsilon \max_\ell \left\{ \chi_{j,\ell} \, \left[\tau_{j,\ell} \, \epsilon_\ell \right]^{1-\gamma} \right\}, & \text{for } 0 \leq \gamma < 1, \\ E_\epsilon \max_\ell \left\{ \ln \left(\tau_{j,\ell} \epsilon_\ell \right) \right\}, & \text{for } \gamma = 1. \\ E_\epsilon \min_\ell \left\{ \left[\chi_{j,\ell} \, \tau_{j,\ell} \, \epsilon_\ell \right]^{1-\gamma} \right\}, & \text{for } \gamma > 1. \end{cases}$$

Also, conditioning on the relevant definition of Φ_j for each γ, we define

$$\bar{\Phi} = \max_j \Phi_j.$$

The following lemma shows that if the average labor market opportunities available to workers are bounded, as summarized by bounds on $\bar{\Phi}$, then, we can guarantee that the dynamic programming problem (6) has a unique and well-defined solution.

Lemma 1 *Let $w \in \mathbb{R}_+^J$ be the vector of unitary wages across all occupations J. Assume that preferences are characterized by a CRRA $\gamma \geq 0$ and that the costs $\tau_{j,\ell}$ and $\chi_{j,\ell}$ and labor market opportunity shocks ϵ satisfy the assumptions above. Then:* ***(a)*** *for all $0 < \gamma \neq 1$, if $\beta \bar{\Phi} < 1$, then there exists a unique, finite $v \in \mathbb{R}^J$ that solves $v^j = \Upsilon^j \frac{\left(w^j\right)^{1-\gamma}}{1-\gamma} + \beta E_\epsilon \max_\ell \left[\chi_{j,\ell} \left\{ \left[\tau_{j,\ell} \, \epsilon^\ell \right]^{1-\gamma} v^\ell \right\} \right]$ for all j. Moreover, if $\gamma < 1$, the fixed point v is positive ($v \in \mathbb{R}_+^J$) and if $\gamma > 1$, the fixed point v is negative ($v \in \mathbb{R}_-^J$).* ***(b)*** *For the special case of log preferences, $\gamma = 1$, if $-\infty < \Phi_j < +\infty$, $\forall j$, and $\beta < 1$, then, there exists a unique, finite $v \in \mathbb{R}^J$ such that $v^j = \Upsilon^j + \ln w^j + \beta E_\epsilon \left[\max_\ell \left\{ v^\ell + \frac{\ln\left(\tau_{j,\ell} \epsilon^\ell\right)}{1-\beta} \right\} \right]$ for all j.*

Appendix A contains the proofs for this and all other analytic results in the paper.

This lemma only verifies existence and uniqueness of the conditional expectations v^j, while the realization ϵ of the labor market opportunities determines the actual realized value $v(j, \epsilon)$. In what follows we impose additional structure so we can characterize the behavior of $v(j, \epsilon)$ and the optimal occupation choices derived for the solution to the dynamic programming problem of workers. To this end, we add the assumption that each element in the vector of labor market opportunities ϵ is distributed according to an extreme value distribution. Specifically, we impose that in each period, *the labor market opportunity ϵ^ℓ shocks for each labor market ℓ, are each independently distributed according to a Frechet distribution with scale parameter $\lambda_\ell > 0$, and curvature $\alpha > 1$.* Notice that the curvature parameter is the same for all occupations but the scale parameters can vary across occupations.

Having impossed a Frechet distribution for ϵ, define for all pairs $j, \ell \in J \times J$,

$$\Omega_{j\ell} = \begin{cases} \chi_{j,\ell} \, \tau_{j\ell}^{(1-\gamma)} v^\ell, & \text{for } \gamma \neq 1, \\ \\ \frac{\ln \tau_{j\ell}}{1-\beta} + v^\ell, & \text{for } \gamma = 1. \end{cases} \tag{7}$$

Then, the normalized BE can be succinctly rewritten as

$$v\left(j,\epsilon\right)=\begin{cases} u^{j}+\beta\max_{\ell}\left\{\Omega_{j\ell}\ \left(\epsilon^{\ell}\right)^{1-\gamma}\right\}, & \text{for } \gamma\neq 1,\\ \\ u^{j}+\beta\max_{\ell}\left\{\Omega_{j\ell}+\frac{\ln\left(\epsilon^{\ell}\right)}{1-\beta}\right\}, & \text{for } \gamma=1.\end{cases} \tag{8}$$

We now provide a simple result that indicates that given any admissible vector $v\in\mathbb{R}^{J}$, regardless of whether it is or not the fixed point of the BE (6), the resulting random variable $v\left(j,\epsilon\right)$ is closely related to one of the extreme value distributions (scaled and displaced): *(a)* if $0\leq\gamma<1$, then $v\left(j,\epsilon\right)$ is related to a *Frechet* with curvature parameter $\alpha/\left(1-\gamma\right)$; *(b)* if $\gamma=1$, then $v\left(j,\epsilon\right)$ is related to a *Gumbel* with shape parameter $1/\alpha$; *(c)* if $\gamma>1$, then $v\left(j,\epsilon\right)$ is related to a *Weibull* with curvature parameter $\alpha/\left(\gamma-1\right)$.

Lemma 2 ***Derived Distributions.*** *Let ϵ^{ℓ} be a random variable distributed Frechet with scale parameter $\lambda_{\ell}>0$ and curvature $\alpha>1$, i.e. its c.d.f. is $F_{\epsilon}\left(\epsilon\right)=e^{-\left(\frac{\epsilon}{\lambda_{\ell}}\right)^{-\alpha}}$. Define:*

$$x_{\ell}\equiv\begin{cases}\left(\epsilon^{\ell}\right)^{1-\gamma} & \textit{for } 0\leq\gamma\neq 1\\ \ln\left(\epsilon^{\ell}\right) & \textit{for } \gamma=1.\end{cases}$$

Then x_{ℓ} is distributed as follows:

$$x_{\ell}\sim\begin{cases}\textit{Frechet}\left(\frac{\alpha}{1-\gamma},\left(\lambda_{\ell}\right)^{1-\gamma}\right) & \textit{for } 0\leq\gamma<1,\\ \textit{Gumbel}\left(\ \frac{1}{\alpha},\ \ln\left(\lambda_{\ell}\right)\ \right) & \textit{for } \gamma=1,\\ \textit{Weibull}\left(\frac{\alpha}{\gamma-1},\ \left(\lambda_{\ell}\right)^{\gamma-1}\right) & \textit{for } \gamma>1.\end{cases}$$

It follows that the terms in curly brackets in (8), with the product of Ω and the transformation of ϵ with respect to the CRRA parameter, will follow one these distributions.[15]

We now complete the characterization of the worker's problem, under the assumption that all the entries of ϵ are independently Frechet distributed. The following theorem provides a simple, sharp characterization for the fixed point problem v^{j} that solves (6) and for the optimal occupations decision of workers.

Theorem 1 ***Individual Problems.*** *Assume for all $\ell=1,...J$, the shocks ϵ_{ℓ} are independently distributed Frechet with shape $\alpha>1$ and scales $\lambda_{\ell}>0$. Assume also that all w^{ℓ} are strictly positive and that either (i) $\gamma\neq 1$ and $\beta\bar{\Phi}<1$ or (ii) $\gamma=1$, $\beta<1$ and $-\infty<\Phi_{j}<+\infty$ for all j. Then:*
(i) *If $0\leq\gamma<1$, the expected values v^{j} for $j=1,...,J$ solve the fixed point problem*

$$v^{j}=\Upsilon^{j}\frac{\left(w^{j}\right)^{1-\gamma}}{1-\gamma}+\beta\,\Gamma\left(1-\frac{1-\gamma}{\alpha}\right)\left[\sum_{\ell=1}^{J}\ \left(\chi_{j,k}v^{\ell}\right)^{\frac{\alpha}{1-\gamma}}\ \left(\tau_{j\ell}\lambda_{\ell}\right)^{\alpha}\right]^{\frac{1-\gamma}{\alpha}}.$$

[15] Note also that Lemma 2 allows us to characterize the value of Υ^{j} given the parameters of the distribution of ϵ^{j}.

A finite solution $v \in \mathbb{R}_+^J$ for this BE exists and is unique. Moreover, the proportion of workers switching from occupation j to occupation ℓ at the end of the period is given by:

$$\mu\left(j,\ell\right) = \frac{\left[\lambda_\ell \tau_{j\ell}\left(\chi_{j,\ell}\, v^\ell\right)^{\frac{1}{1-\gamma}}\right]^\alpha}{\sum_{k=1}^{J}\left[\lambda_k \tau_{jk}\left(\chi_{j,\ell}\, v^k\right)^{\frac{1}{1-\gamma}}\right]^\alpha}.$$

(ii) *If $\gamma = 1$ the expected values v^j for $j = 1,...,J$, solve the fixed point problem*

$$v^j = \Upsilon^j + \log(w^j) + \frac{\beta}{\alpha\left(1-\beta\right)}\log\left[\sum_{\ell=1}^{J}\exp\left(\alpha\left(1-\beta\right)v^\ell + \alpha\log(\tau_{j\ell}) + \alpha\log(\lambda_\ell) + \alpha\,\kappa\right)\right],$$

where κ is Euler's constant. A solution $v \in \left[\underline{v}, \bar{v}\right]^J$ for this BE exists and is unique. Moreover, the proportion of workers that switch from occupation j to occupation ℓ is given by:

$$\mu\left(j,\ell\right) = \frac{\exp\left(\alpha\left(1-\beta\right)v^\ell + \alpha\log(\tau_{j\ell}) + \alpha\log(\lambda_\ell) + \alpha\,\kappa\right)}{\sum_{k=1}^{J}\exp\left(\alpha\left(1-\beta\right)v^k + \alpha\log(\tau_{jk}) + \alpha\log(\lambda_k) + \alpha\,\kappa\right)}.$$

(iii) *If $\gamma > 1$, the expected values v^j for $j = 1,...,J$ solve the fixed point problem*

$$v^j = \Upsilon^j\frac{\left(w^j\right)^{1-\gamma}}{1-\gamma} - \beta\,\Gamma\left(1 - \frac{1-\gamma}{\alpha}\right)\left[\sum_{\ell=1}^{J}\left(-\chi_{j,\ell}\, v^\ell\right)^{\frac{\alpha}{1-\gamma}}\left(\tau_{j\ell}\lambda_\ell\right)^\alpha\right]^{\frac{1-\gamma}{\alpha}}.$$

A solution $v \in \left[\underline{v}, 0\right]^J$ for this BE exists and is unique. Moreover, the proportion of workers switching from occupation j to occupation ℓ at the end of the period is given by:

$$\mu\left(j,\ell\right) = \frac{\left[\lambda_\ell \tau_{j\ell}\left(-\chi_{j,\ell}\, v^\ell\right)^{\frac{1}{1-\gamma}}\right]^\alpha}{\sum_{k=1}^{J}\left[\lambda_k \tau_{jk}\left(-\chi_{j,k}\, v^k\right)^{\frac{1}{1-\gamma}}\right]^\alpha}.$$

3.4 Implied Distributions of Workers and Human Capital

We now describe how the occupation choices of each worker shape up the limiting behavior of the cross-section distribution of workers and aggregate human capital (and earnings) across the J occupations in this economy.

Distribution of Workers Across Occupations. Notice that the homogeneity of the value function implies that the transitions $\mu^{j,\ell}$ are independent of the worker's level of human capital, h, i.e. his absolute advantage. Let $\theta_t = \left[\theta_t^1, ..., \theta_t^J\right]$ denote the $1 \times J$ vector indicating the mass of workers in each of the occupations $j = 1, 2, ..., J$ at time t. Since in this section we are taking the vector of wages w as time invariant the transition matrix μ is also time invariant. Therefore, the evolution of θ is described by following equation,

$$\theta_{t+1} = \theta_t \, \mu.$$

Under the assumptions that $\tau_{j\ell} > 0$, every entry of the stochastic matrix μ is positive, i.e. for all j, ℓ, $\mu^{j\ell} > 0$. Under this basic mixing condition, standard results for Markov chains (e.g. Theorem 11.2. in Stokey, Lucas Jr, and Prescott (1989)) it follows that there exists a unique invariant distribution

$$\theta_\infty = \theta_\infty \, \mu, \tag{9}$$

and that from any initial distribution θ_0, the employment distribution across occupations will converge to it, i.e.: $\lim_{t\to\infty} \theta_0 \mu^t = \theta_\infty$.

Distribution of Aggregate Human Capital Across Occupations. Given that the individual labor market opportunities or productivity shocks for all workers are distributed Frechet, a continuous distribution with full support in the positive reals, then the aggregate human capital assigned to occupation j is given by,

$$H_t^j = \theta_t^j \int_0^\infty h \phi_t^j(dh),$$

where $\phi_t^j(\cdot)$ denotes the positive measure that describes the distribution of human capital levels h across the workers in occupation j in period t.

Characterizing the evolution of H_t^j over time suffices to determine the general equilibrium of the economy as we discuss in the following section. Towards that end, we first characterize the conditional expectation of the shocks ϵ^ℓ for those workers that switch from any occupation j to any occupation ℓ:

Lemma 3 *For all non-negative $\gamma \neq 1$, the expectation of the labor market opportunity shock ϵ_ℓ of workers switching from j to ℓ is given by the following equation:*

$$E\left[\epsilon_\ell | \Omega_{j\ell} \, \epsilon_\ell^{1-\gamma} = \max_k \left\{ \Omega_{j\ell} \epsilon_k^{1-\gamma} \right\}\right] = \Gamma\left(1 - \frac{1}{\alpha}\right) \lambda_\ell \left[\mu^{j\ell}\right]^{-\frac{1}{\alpha}}, \tag{10}$$

where $\mu^{j\ell}$ is the corresponding occupation switching probabilities as derived in Theorem 1.

A worker with human capital h in occupation j will switch to occupation ℓ at the end of the period with probability $\mu^{j\ell}$, bringing an average $\Gamma\left(1 - \frac{1}{\alpha}\right) \tau_{j\ell} \, \lambda_\ell \left[\mu^{j\ell}\right]^{-\frac{1}{\alpha}} h$ of human capital skills to that occupation. Define $\mathcal{M}$ to be the transition matrix of aggregate human capital, with j, ℓ element defined as:

$$\mathcal{M}^{j,\ell} = \Gamma\left(1 - \frac{1}{\alpha}\right) \tau_{j\ell} \, \lambda_\ell \left[\mu^{j\ell}\right]^{1-\frac{1}{\alpha}}.$$

The matrix $\mathcal{M}$ is time invariant when wages are constant over time. The linearity in h allows an easy aggregation of human capital in each occupation and also to characterize the law of motion for aggregate human capital. Let $H_t = \left[H_t^1, H_t^2, ..., H_t^J\right]$ be the $1 \times J$ vector of aggregate human capital across all occupations j in period t. Then, for time $t+1$, that vector evolves according to

$$H_{t+1} = H_t \, \mathcal{M}.$$

It is worth remarking that we can characterize the evolution of the average (or total) skills of workers across the different occupations, without having to solve for the within-occupation cross-section distribution of skills and earnings. This result is useful since we can easily solve for the aggregate supply of efficient units of labor in each occupation at each t. The matrix $\mathcal{M}$ is strictly positive, i.e. $\mathcal{M}^{j\ell} > 0$. Then, from the Perron-Frobenius theorem, the largest eigenvalue of $\mathcal{M}$ is always simple (multiplicity one), real, and positive.[16] Moreover, the associated eigenvector to this so-called Perron root, which we denote by G_H, has all its coordinates, H^j, $j = 1, ..., J$, strictly positive. Moreover, in the limit, the behavior of all H_t^j will converge to

$$H_{t+1}^j = G_H H_t^j,$$

for all $j = 1, ...J$. This is precisely the definition of a balanced-growth path (BGP) for the vector of aggregate human capital $\{H_t\}_{t=0}^{\infty}$. Notice that the model can naturally generate a Perron eigenvalue greater than one $G_H > 1$, i.e. sustained growth of the human capital of the workers, even if the unitary wages w^j and the cross-section distribution of workers θ_∞ remains constant, and even if the average realization ϵ^ℓ in each occupation is equal to (or even lower than) one. The engine of growth of the individual worker's earnings is that workers continuously select the most favorable labor market opportunities.[17]

We summarize the results for the implied population dynamics of workers and human capital aggregates, $\{\theta_t, H_t\}_{t=0}^{\infty}$ in the following proposition.

Proposition 1 *Assume that the unitary wage vector is strictly positive, $w \in \mathbb{R}_{++}^J$, and that the conditions for Theorem 1 hold. Then:* ***(a)*** *There exists a unique invariant distribution of workers, i.e., $\theta_\infty = \theta_\infty \, \mu$, with $\theta_\infty^j > 0$ for all j and $\sum_{j=1}^J \theta_\infty^j = 1$. Moreover, the sequence $\{\theta_t\}_{t=0}^{\infty}$ induced by (9) converges to θ_∞ from any initial distribution θ_0.* ***(b)*** *There is a unique BGP of aggregate human capital across occupations, $H_t^j / H_t^1 = h^j$ for all j, where h^j is equal to the ratios of the j^{th} coordinate to the first coordinate of the Perron eigenvector. Moreover, the economy converges to $H_{t+1} = G_H \cdot H_t$ from any initial vector $H_0 \in \mathbb{R}_+^J$.*

The problem of the worker presented so far can be easily extended to capture worker heterogeneity along permanent characteristics (gender, race, formal education) as well as age. As shown

[16] See for example, Gantmacher (1959), Theorems 1 and 2 of Ch.XIII, Vol. II.

[17] This result is reminiscent of the mechanism in the models by Luttmer (2007) and Lucas and Moll (2014) in which selection on favorable realization of idiosyncratic shocks endogenously generates growth at the aggregate level. Note however that in our model it is possible for a single worker to get a realization of shocks ϵ below one for all components, which implies that human capital will decrease for this individual if $\tau \leq 1$.

in Appendix D, the setting can be quite flexible in allowing differences in group specific parameters ($\lambda_{\ell}^{\text{group}}$, $\tau_{j\ell}^{\text{group}}$, $\chi_{j\ell}^{\text{group}}$), thus allowing differentiating between the human capital accumulation that arises from labor market experience from other factors that affect the human capital of workers. Extending the model for age differences would capture differences in the horizon of workers and their dynamic valuation of switching occupations.

In the Section 4, we embed the workers' problem into a production economy, and extend the results derived here to characterize the general equilibrium of such an environment. Then, in Section 5, we use the model to quantitatively examine the dynamics of occupation choices, income inequality and growth in the U.S. economy. To this end, we craft our model within a simply demographic structure in which new workers enter very period and some active workers retire.

3.5 Cohorts

We embed our model into a simple demographic structure with workers of multiple ages. Concretely, we employ a simple perpetual youth model, in which a constant mass $0 < \delta < 1$ of new workers enter every period, and every period, the same fraction of previously active workers die and leave the economy. Thus, the discount factor in our previous analysis encompasses both a pure time discount factor and a survival probability $(1 - \delta)$. However, we now need to specify the decisions of new workers, and to account for the overall population of workers in each period, which is composed by cohorts of many different ages.

3.5.1 New Workers

In the period before entering the labor market, all new workers have one unit of human capital, $h^{new} = 1$ and face an occupation choice similar to that of the active workers. However, new workers are not attached to any particular occupation. Their choices will therefore be determined by vectors $\tau^0 = \left[\tau^{0,1}, \tau^{0,2}, \ldots, \tau^{0,J}\right]$ and $\chi^0 = \left[\chi^{0,1}, \chi^{0,2}, \ldots, \chi^{0,J}\right]$, which affect the amount of human capital and the employment decisions of new workers in each occupation.[18]

Under these assumptions, the vectors of employment shares and initial human capital of new workers across occupations are given by

$$\theta^0(j) = \frac{\left[\lambda_j \tau^{0,j} \left(-v^j\right)^{\frac{1}{1-\gamma}}\right]^{\alpha}}{\sum_{k=1}^{J} \left[\lambda_k \tau^{0,k} \left(-v^k\right)^{\frac{1}{1-\gamma}}\right]^{\alpha}},$$

and

$$H^0(j) = \Gamma\left(1 - \frac{1}{\alpha}\right) \tau^{0,j} \, \lambda_j \left[\theta^{0,j}\right]^{1-\frac{1}{\alpha}}.$$

[18] As in the previous case, these two different set of parameters, τ^0 and χ^0 provides the model with the flexibility to match both the occupational choices of entering cohorts as well as the differences in average earnings (human capital) across occupations for the entering cohort.

Both expressions are similar to those of active workers. They highlight how entering workers choose their initial occupation according to the values v^j and their average initial human capital is determined by their optimal job selection.

3.5.2 Stationary Distributions of Workers and Human Capital

Then, we compute the values of v, the transition matrices μ, $\mathcal{M}$, and the entrants employment and human capitals θ^0, H^0. Then, for any age $s = 1, 2, ...,$ the employment shares will follow $\theta^s = \mu\,\theta^{s-1}$. Out of these, only a fraction $(1-\delta)$ survives. As for human capital, the total mass of human capital for the next period of each cohort of workers is given by $H^s = (1-\delta)\,\mathcal{M}H^{s-1}$. For the economy as a whole, the steady state employment shares across all cohorts is given by

$$\theta = \delta\,\theta^0\,[I - (1-\delta)\,\mu]^{-1}\,.$$

which always exists, since μ is a stochastic matrix and its highest eigenvalue is 1. Similarly, the steady state J-vector of aggregate human capital levels accross all cohorts would be given by

$$H = \delta\,H^0\,[I - (1-\delta)\,\mathcal{M}]^{-1}\,.$$

It is interesting and important to highlight that H might not be well defined. Indeed, if the Perron root of $\mathcal{M}$ is higher than $(1-\delta)^{-1}$, then the average growth of the surviving workers more than compensates the rate at which they die, the aggregate human capital of each cohort grows over time, i.e. $H^s > (1-\delta)\,\mathcal{M}H^{s-1}$, and the country's human capital vector $H_t = \sum_{s\leq t} H^s$ would grow without bounds over time, eventually, at the rate $(1-\delta)\,G_H$. We now illustrate the behavior of the economy, highlighting the impact of relative unitary wages in the determination of the aggregates. We re-state the implications of our dynamic Roy model with cohorts for the implied dynamics of workers and human capital aggregates, for the population as a whole and for each cohort.

Proposition 1 (Cohorts) *Assume that the unitary wage vector is strictly positive, $w \in \mathbb{R}^J_{++}$, and that the conditions for Theorem 1 hold. Then:* ***(a)*** *There exists a unique invariant distribution of all workers,*

$$\theta = \delta\,\theta^0\,[I - (1-\delta)\,\mu]^{-1}\,.$$

where the initial distribution θ_0 is induced by the optimal occupation choices of young workers. ***(b)*** *Generically, there is either:* ***(i)*** *a unique stationary stock of human capital across occupations if $G_H(1-\delta) < 1$. Such a stock is given by*

$$H = \delta\,H^0\,[I - (1-\delta)\,\mathcal{M}]^{-1}\,,$$

where the initial distribution H_0 is induced by the optimal occupation choices of young workers;

*or **(ii)** a unique BGP of aggregate human capital across if $G_H(1-\delta) > 1$. If so, all entries in the vector of human capitals asymptotically grow at the rate $G_H(1-\delta)$ and settle in the Perron (dominant) eigenvector, i.e. $H_t^j/H_t^1 = h^j$ for all j, where h^j is equal to the ratios of the j^{th} coordinate to the first coordinate of the Perron eigenvector of $\mathcal{M}$.*

3.6 A Numerical Illustration

In this section, we illustrate the key mechanisms of the model. We center our discussion around two main themes. First, we highlight the importance of dynamic occupation choices, not only for the aggregate allocation of labor across jobs types, but also for the long run income level and its distribution across the different workers. Second, we use these numerical illustrations to advance the analysis of the impact of task biased technical change that affect occupations asymmetrically.

For these numerical illustrations, we aim to mimic an economy in which occupations differ along two dimensions: **(a)** jobs that are either manual (M) or cognitive (C), and **(b)** jobs that are either routine (R) or non-routine (N). Thus, we consider an economy with four occupations, $J = 4$, that differ in both their unitary skill prices, w_j, and their implied dynamic path for the human capital of the workers, as delineated by the transferability of skills $\tau(j, \cdot)$.

We set the vector of unitary skill prices $\mathbf{w}$ and the transferability matrix $\tau[j, \ell]$ to be:

$$\mathbf{w} = \begin{bmatrix} w_{\mathbf{MR}} \\ w_{\mathbf{MN}} \\ w_{\mathbf{CR}} \\ w_{\mathbf{CN}} \end{bmatrix} = \begin{bmatrix} 1.25/\mathbf{tbtc} \\ 1 \\ 1.25/\mathbf{tbtc} \\ 1 \end{bmatrix}, \text{ and } \tau[j, \ell] = \begin{bmatrix} 1.000 & \tau_{RN} & \tau_{MC} & \tau_{MC} \cdot \tau_{RN} \\ \tau_{NR} & 1.015 & \tau_{MC} \cdot \tau_{NR} & \tau_{MC} \\ \tau_{CM} & \tau_{CM} \cdot \tau_{RN} & 1.015 & \tau_{RN} \\ \tau_{CM} \cdot \tau_{NR} & \tau_{CM} & \tau_{NR} & 1.050 \end{bmatrix},$$

where *tbtc* is a 'task-biased technological change' shifter that *reduces the productivity of routine occupations*. We use this shifter to compare different economies as explained shortly.

For the transferability of skills, we impose a simple structure whereby the costs of switching occupations depend on whether the current job differs from the new job in one or two dimensions and also the direction of change. Specifically, we set $\tau_{CM} = 0.95$, $\tau_{NR} = 0.95$, $\tau_{MC} = 0.7$, and $\tau_{RN} = 0.6$. With these numbers, a worker that switches from a cognitive to a manual job, but otherwise maintains the routine or non-routine nature of his occupation, would be able to transfer, on average 95% of his skills. If on top of that, the worker is also switching from a non-routine to a routine job, he will only be able to transfer on average $90.25\% = \tau_{CM} \cdot \tau_{NR}$ of his existing skills. Moreover, this parametrization imposes that it is more difficult to switch from routine to non-routine and from manual to cognitive than the respective moves in the opposite directions. Finally, we parametrize the diagonal entries of the matrix so that staying in manual-routine jobs implies average zero growth in human capital, staying in manual-non routine or cognitive-routine jobs lead to low growth, 1.5% per period. Cognitive-non routine jobs are not better paid in terms of unitary wages, but they lead to a much higher growth rate, 5% per period, for the worker's human capital and earnings.

Having thus parametrized the matrix τ, we normalize the Frechet distribution of labor market shocks ϵ_j to be $\lambda = [1,\, 1,\, 1,\, 1]' / \Gamma(1 - 1/\alpha)$, where we set the curvature of those distributions to be $\alpha = 15$. We assume for the discount factor, $\beta = 0.95$, a standard value for an annual model. We set CRRA, $\gamma = 2$, also standard value in quantitative macro. In these illustrations, we set $\delta = 0.04$. Finally, for the new workers, we set the transferability vector to be $\tau^0 = [1,\, 1,\, 1,\, 0.65]$. In this way, new workers can more easily enter occupations that are either manual or routine, while cognitive non-routine jobs are more difficult to find right off the bat as their initial occupation. We emphasize that these parameter choices are chosen just to illustrate the mechanics of our model. Our quantitative exercises are in Section 5.

We use this parametrization of our model to compare economies with values of tbtc that range from 1 to 1.25. A higher value of tbtc can be seen as task-biased technological change because it reduces the price of routine labor, both in manual and cognitive occupations. For clarity, we keep the price of both non-routine occupations constant, i.e. tbtc does not lead to an increase in their price. Under these circumstances, we highlight a key mechanism of our model: that the reallocation of workers from routine to non-routine can lead the economy to sustain much higher levels of aggregate human capital and per capita income. This mechanism would be even stronger if tbtc also increases the productivity of non-routine occupations.

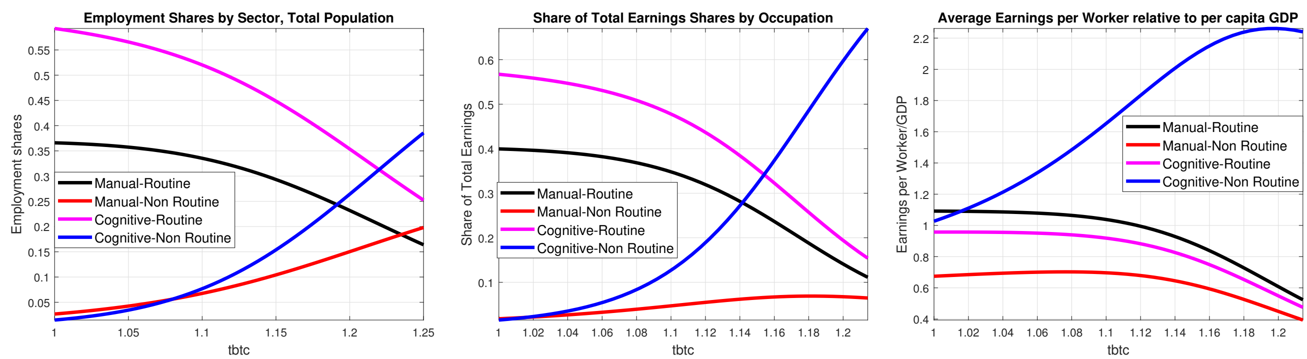
We divide our numerical illustrations in three parts. First, we compare the steady-states associated with different levels of tbtc. Second, we examine the transition dynamics of an economy that permanently moves to a higher task-biased technological level. Third, we examine the impact of mobility costs, both across steady-states or BGPs and for transitional dynamics.

Overall, these exercises show that the endogenous response in the aggregate accumulation of human capital can lead to substantial differences across steady-states. Indeed, the response in aggregate human capital accumulation can be so strong that higher levels of tbtc lead to a regime change, whereby the economy switches from having a steady state to exhibit endogenous sustained growth. For the transition dynamics, we show that the model generates non-monotone and a very slow transition that, for many periods, can be easily confused with sustained growth. Finally, we find that the costs of moving human capital across occupations have a profound impact on the aggregate dynamics, not only across steady-states and BGPs, but also for the transition from a steady state with low tbtc to another with a higher tbtc.

3.6.1 Cross Steady-States Comparisons

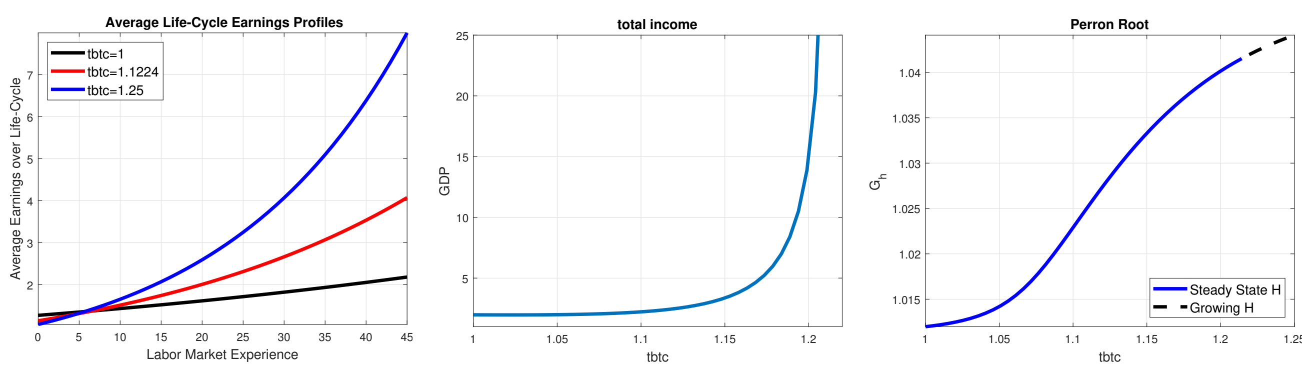
We start by showing that task-biased technical change in our model economy can easily replicate the patterns of job market polarization and between-groups earnings inequality discussed in Section 2. We compute the implied steady states associated with the vectors of unitary wages $\mathbf{w}$ for values of tbtc ranging from 1 to 1.25. For low tbtc levels, the wages of routine occupations are high and as shown in panel (a) of Figure 2, their employment shares in those occupations (black and magenta) would almost completely dominate the economy. For higher levels of tbtc, the unitary wages of non-routine catch up and their implied faster accumulation of skills would attract workers to those occupations. In this example, for tbtc close to 1.25, non-routine occupations become dominant. This pattern mimics the labor market polarization discussed above, and is even stronger when we look at the human capital as shown in panel (b) of Figure 2 with the share of total earnings of each occupations. Finally, panel (c) of the figure shows that a strong widening gap between the average earnings of workers in cognitive non-routine occupations relative to all the others. Here, we compute the values $w_j H_j^T / \theta_j^T$ relative to the average GDP. It shows that while in an economy with tbtc=1, cognitive non-routine workers would earn on average the same earnings as manual-routine and cognitive-workers, once those workers are in an economy with tbtc=1.25 they can earn up to 5 times more than the workers in the other occupations. Obviously, these numbers are not driven by the pure price impact of tbtc because it is much smaller, but instead it is driven by the sustained human capital accumulation of workers in cognitive/non-routine occupations.

Figure 2: Employment & Earnings: BGPs across Task-Biased Technological Levels



(a) Total employment shares. (b) Output shares. (c) Average Earnings w.r.t. GDP.

Figure 3: Employment & Earnings: BGPs across Task-Biased Technological Levels



(a) Life-Cycle Earnings Growth (b) Aggregate GDP (c) Perron root of $\mathcal{M}$

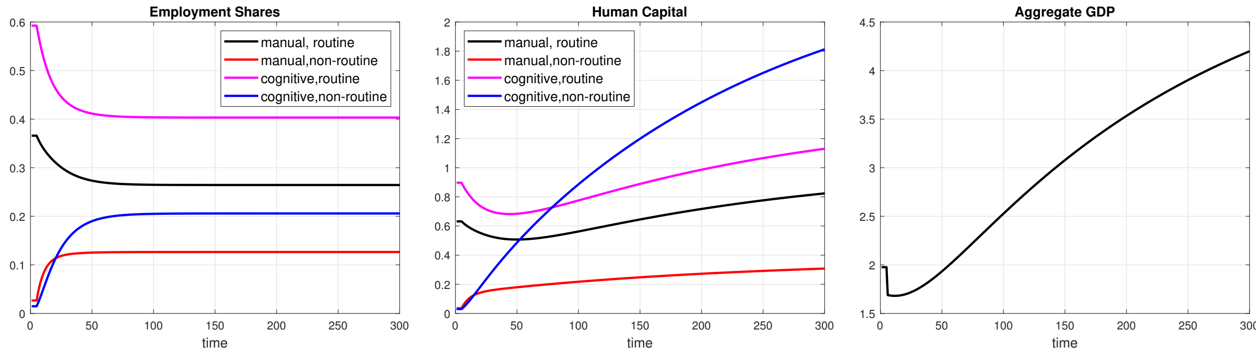
The reallocation of workers across occupations can have an enormous impact for the behavior of aggregate output. Panel (a) of Figure 3 shows the average life-cycle earnings profile of workers in economies with different values of tbtc. With tbtc=1 (black line), most workers would opt for routine occupations and, on average, accumulate human capital very slowly. With tbtc=1.125 (red-line), young workers would earn less in the beginning, but their life-cycle profiles are much steeper. Instead of just doubling their income after 45 years of labor market experience, their earnings would be more than seven times higher. Much steeper profiles accrue when tbtc=1.25 (blue line). Hence, with differences in tbtc levels, our model naturally replicates the patterns unveiled by Lagakos et al. (2018) across countries with different levels of development. At the aggregate level, these differences in the accumulation of human capital can lead to very large income differences as shown in panel (c) of Figure 3. A change from tbtc=1 to tbtc=1.2 would be associated with (eventually) multiplying the income of the country by more than 7 times. In fact, this simple mechanism is so strong that it leads to a bifurcation in the behavior of income. As shown in the panel (b), GDP has a vertical asymptote around 1.22, and aggregate output can be arbitrary high for values of tbtc close enough to it. As explained above and illustrated in panel (c) of Figure 3, the reallocation of workers in the economy leads to a change in the overall growth in their human capital which can lead to a bifurcation in the aggregate economy. This is captured by the implied Perron Root of the matrix $\mathcal{M}$ for different values of tbtc. Here, the solid-blue line denotes the case when $G_H < (1-\delta)^{-1}$, and the economy settles in a steady state. The dashed-black curve capturers the cases in which the above condition is not met and then the aggregate human capital of the country would keep growing and no steady-state could exist. Under the parameter values of the exercise, the point of bifurcation is when tbtc has a value of 1.22. Above that bifurcation, aggregate output of the country would grow at the endogenous rate $G_H(1-\delta)$ every period.

3.6.2 Transition Dynamics

We now explore the aggregate transitional dynamics of an economy that experiences a once-and-for-all task biased technical change. Specifically, we assume that the economy is initially at **tbtc=**1 and then, unexpectedly but permanently at $t = 6$, it shifts to **tbtc=**1.2. Then after, the wages of routine occupations fall permanently, and all workers, including the new and future ones, will take those new wages as given. Recall that for clarity, non-routine unitary wages remain constant.

Panels (a) and (b) of Figure 4 display the response, over time, in the employment shares and the human capital allocated to the four occupations. As expected, workers in the routine occupations start moving away from them towards the non-routine occupations. This is a long and persistent process because of the substantial costs of existing workers when switching occupations. Yet, much of the reallocation of workers happens in the first twenty years, and after fifty periods it is almost done. The transition for human capital levels across occupations requires much more time. As shown in the figure, from being a negligible fraction, the human capital in cognitive non-routine occupations becomes the dominant one. Aggregate human capital needs a long time to settle into the new steady state level. Even after three hundred period the economy is far from

Figure 4: Aggregate Transition: from tbtc=1 to tbtc=1.2



(a) Employment Shares (b) Human Capital (c) Aggregate Income

the new balanced growth path.

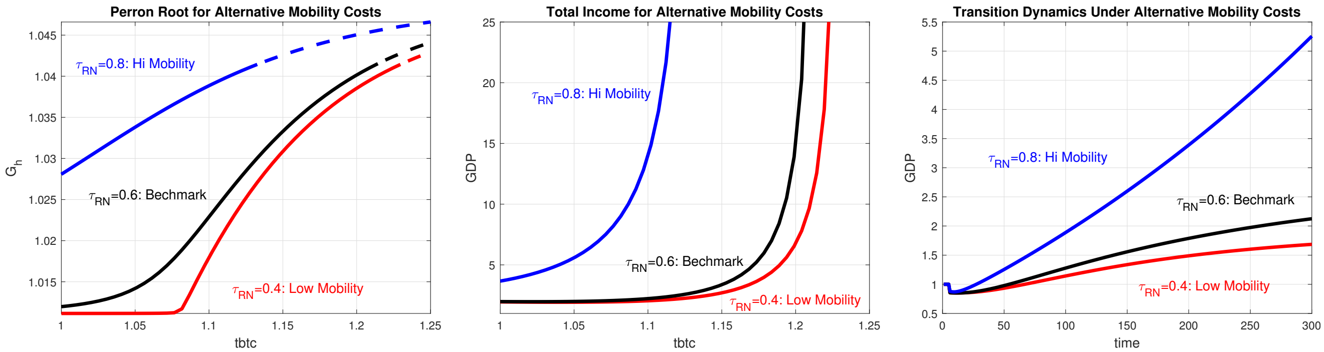
Most interesting, this once-and-for all change leads to a non-monotone and long-protracted transition. To be sure, since the fall in the wages of routine workers are without a counterpart on those of non-routine, they necessarily lead to an initial collapse in aggregate output. After approximately ten years of stagnation, output starts to grow, but it needs about fifty periods to reach again the initial levels. From then on, the economy will keep growing for many periods. In this example, it requires more than 600 periods for the economy to really get close to the new steady state.

3.6.3 Aggregate Consequences of Occupation Mobility Costs

Finally, in this section, we examine the aggregate implications of occupation mobility costs. To this end, we compare the previous economy (benchmark) with two economies in opposite sides of mobility costs. For concreteness, we focus on the cost of moving from routine to non-routine jobs. Our benchmark with $\tau_{RN} = 0.6$ is compared with:**(a) an economy with lower mobility costs** with $\tau_{RN} = 0.8$, and with **(b) an economy with higher mobility costs** with $\tau_{RN} = 0.4$. Keeping all other parameters the same, we compare the economies both across steady states/BGPs for different tbtc levels and for the same transitional dynamics as above.

Panels (a) and (b) of Figure 5 present the results for the cross steady states/BGPs. Panel (a) clearly shows that mobility costs have a very strong detrimental impact on the growth rate in the workers' earnings over the life-cycle, specially for lower levels of tbtc. Interestingly, when the costs of mobility are high (red line), advances in tbtc can be neutralized and have no impact on the life-cycle profiles and the overall aggregate economy. For higher levels of tbtc, the differences in mobility costs can be quite large. The implied growth rate of earnings is much higher with high mobility (high τ_{RN}, blue line.) Panel (b) shows that the effect on aggregate output can be quite high. For instance, with tbtc=1.1, the output with high mobility can be up to fifteen times the output in the other two economies. For higher levels of tbtc, the implications can be even

Figure 5: BGPs and Transitions: Task-Biased Technological Levels and Mobility Costs



(a) Perron-Root of $\mathcal{M}$.

(b) Output Levels.

(c) GDP transition dynamics

more dramatic. Notice that the bifurcation point between steady states and BGPs takes place at a much lower level in the high mobility economy than in the benchmark. For the low mobility economy, the bifurcation point is at a higher level than in the benchmark.

Finally, panel (c) of Figure 5 presents the transition dynamics for the three economies after the same shock as before from tbtc=1 to tbtc=1.2. To facilitate the comparison, aggregate outputs have been normalized to the initial level. The qualitative features of the aggregate responses are similar, but the actual impact over time depends on the level of mobility. Economies with lower mobility settle sooner partly because of a smaller final impact. Notably, for economies with high mobility, the change implies shifting from a stagnant economy to one that will grow forever after.

4 The General Equilibrium Model

We now set up our general equilibrium environment. First, we specify the production of final goods, which defines the demand for the different types of labor and capital and the production price of final goods. Second, we define competitive equilibria, where the price of goods, labor services and capital clear all markets. Third, we provide a sharp characterization of the intratemporal equilibrium conditions. Finally, we prove the existence of balance growth path (BGP) equilibria.

4.1 The Environment

4.1.1 Production

We consider multiple types of workers and physical capital as factors of production of final goods. Our setting encompasses features of the standard neoclassical model and of recent models of substitution between workers and machines (e.g. Acemoglu and Restrepo (2018)), within a worker-task assignment model (e.g. Costinot and Vogel (2010).) First, as in standard macro models, we allow for some forms of physical capital to operate as a complementary factor of all forms of labor. Second, as in Acemoglu and Restrepo (2018), we also allow for some other forms of physical capital (machines) to compete with workers in the performance of tasks. As Costinot and Vogel (2010), different types of workers must be be assigned across multiple production tasks according to their comparative advantage, which is determined in general equilibrium. The resulting multidimensional production setting allows for technological changes that have a heterogeneous impact on the different types of labor.

Consider an economy with a single final good, which is produced according to a Cobb-Douglas over structures and other forms of physical capital, K_t, and a bundle of tasks, Q_t,

$$Y_t = (K_t)^{\varphi} (Q_t)^{1-\varphi},$$

where $0 < \varphi < 1$.[19] The bundle of tasks Q_t is given by a CES production function defined over many tasks. There are J occupations and in each occupation a continuum of tasks must be performed. The production of aggregate tasks is defined as

$$Q_t = \left[\sum_{j=1}^{J} \left(Q_t^j\right)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}},$$

where Q_t^j is the output of tasks from each occupation j and ρ is the elasticity of substitution across the output from the different occupations. For each j, the output Q_t^j arises from the use of workers and machines performing many different tasks. We take the set of those tasks to be the continuum $[0,1]$. In particular, we assume that

$$Q_t^j = \left(\int_0^1 \left[q_t^j(x)\right]^{\frac{\eta-1}{\eta}} dx \right)^{\frac{\eta}{\eta-1}},$$

where $\eta > 1$ is the elasticity of substitution across the different tasks.

For each $x \in [0,1]$, the output $q_t^j(x)$ is obtained from the use of human capital H_t^j or machines M_t^j in that occupation and task. Extending Acemoglu and Restrepo (2018), we assume that machines and labor are *perfect substitutes* to each other in the production of each task x in occupation j. The production function of $q_t(x)$ is described by

$$q_t^j(x) = z_t^{j,M}(x) M_t^j(x) + z_t^{j,H}(x) H_t^j(x), \tag{11}$$

where $= z_t^{j,M}(x)$ is the productivity of machines performing task x in occupation j, and $z_t^{j,H}(x)$ is the productivity of workers in that occupation performing this task x. Here, $H_t^j(x)$ and $M_t^j(x)$ are total effective units of labor j and machines used in task x.[20]

[19] In this way, our Cobb-Douglas assumption between structures and the other factors of production follows closely Krusell et al. (2000).

[20] The production function of tasks in equation (11) can easily be extended to incorporate several factors of production, like different types of machines (computers, robots) or different types of labor (non-college educated,

For tractability, we assume that for all j, $x \in [0,1]$ and periods t, the productivities of labor j and machines are the products of time-varying components and a time-invariant components. For each j, the time-varying component is common across all $x \in [0,1]$. The task specific components differs across occupations j but is fixed over time. Then, the productivity level workers and machines are, respectively, $z_t^{j,H}(x) = A_t^{j,H}\, z^{j,H}(x)$ and $z_t^{j,M}(x) = A_t^{j,M}\, z^{j,M}(x)$. Furthermore, for each j, we assume that the time-invariant components are distributed i.i.d. Frechet across the different tasks with shape parameter $\nu_j > 1$. For all j, we normalize the scale parameters to to one. In this way, we can use the results in Eaton and Kortum (2002) to further characterize optimal demands of factors of production for the different tasks and the overall production cost of the good as we discuss below.

4.1.2 Capital Owners

We assume that the two forms of physical capital, machines M_t^j across all occupations j and structures and other equipment K_t are owned by a separate set of households. These households, which we call 'capital owners,' have a constant population with measure 1. Capital owners have standard preferences, given by

$$U_t^K = \sum_{s=0}^{\infty} \beta^t \frac{\left(c_t^K\right)^{1-\gamma}}{1-\gamma}, \tag{12}$$

where, for simplicity, we have assumed that the discount factor β and the CRRA γ have the same values as those of the workers, however, we need to assume $\gamma > 0$ for an interior solution on the investment problem.

Both forms of capital follow neoclassical laws of motion:

$$K_{t+1} = \left(1-\delta^K\right) K_t + \xi_t^K I_t^K, \quad \text{and} \tag{13}$$

$$M_{t+1}^j = \left(1-\delta^M\right) M_t^j + \xi_t^M I_t^{M,j}, \;\; \forall j = 1, 2, \ldots J, \tag{14}$$

where $0 < \delta^K, \delta^M < 1$. Both forms of investment I_t^K, $I_t^{M,j}$ are in units of the final good. The strictly positive terms ξ_t^K and ξ_t^M capture investment specific productivities. We assume that the same parameters δ^M, ξ^M govern the law of motion for the machines in all j. Then, in equilibrium, the return to investing in machines must be equalized across all j, and we only need to keep track of a single rental price of machines.

Capital owners rent out machines and structures, taking as given their rental prices, r_t^M, and r_t^K. Capital owners can freely borrow or lend at the gross (real) interest rate R_t/P_t. We denote by B_t the net financial position of the representative capital owner in period t. In terms of financial markets, below we consider two polar cases. First, we consider a small open economy in the interest rate R_t in every period is taken exogenous from international capital markets. Second, we

college educated), each with different comparative advantages in the production of tasks. All of the expressions we derive next easily extend to this more general case.

consider a closed economy equilibrium in which $B_t = 0$ for all periods.

4.2 Competitive Equilibria

We assume all labor, capital and goods markets are perfectly competitive. Taking as given the sequence $\left\{P_t, w_t^j, r_t^K, r_t^M, R_t\right\}_{t=0}^{\infty}$ of goods prices, the unitary skill price for jobs of all types $j = 1, ..., J$ and the rental rate of both forms of capital, firms and households maximize their current profits and expected lifetime utilities, respectively. To formally define competitive equilibrium in this environment, we first define the individual problems of firms and workers and outline the market clearing conditions.

4.2.1 Workers' Optimization and Aggregate Labor Supply

The maximization problem of each of the workers is simply the time-varying extension of the problem characterized in Section 2. For brevity, we only consider here the case of $\gamma > 1$, as the other cases are similar. For every t, j and h, the expected normalized values $\left\{v_t^\ell\right\}_{\ell=1}^{J}$ of an active worker solve the problem recursion:

$$v_t^j = \frac{(w_t^j)^{1-\gamma}}{1-\gamma} - \beta\,\Gamma\left(1 - \frac{1-\gamma}{\alpha}\right)\left[\sum_{\ell=1}^{J}(-\chi_{j,\ell}\,v_{t+1}^\ell)^{\frac{\alpha}{1-\gamma}}\ (\tau_{j\ell}\lambda_\ell)^\alpha\right]^{\frac{1-\gamma}{\alpha}}, \tag{15}$$

and their optimal occupation choices, i.e. transitions from any j to any ℓ are given by

$$\mu_t^{j\ell} = \frac{\left[\lambda_\ell\tau_{j\ell}\left(-\chi_{j,\ell}\,v_{t+1}^\ell\right)^{\frac{1}{1-\gamma}}\right]^\alpha}{\sum_{k=1}^{J}\left[\lambda_k\tau_{jk}\left(-\chi_{j,k}\,v_{t+1}^k\right)^{\frac{1}{1-\gamma}}\right]^\alpha}, \tag{16}$$

where $\left\{v_{t+1}^\ell\right\}_{\ell=1}^{J}$ solves the problem for the subsequent period. Similarly, the transition matrix for aggregate human capital from occupation j to occupation ℓ for the time-varying case is simply

$$\mathcal{M}_t^{j\ell} = \Gamma\left(1 - \frac{1}{\alpha}\right)\tau_{j\ell}\,\lambda_\ell\left[\mu_t^{j\ell}\right]^{1-\frac{1}{\alpha}}. \tag{17}$$

The implied laws of motion for the population of workers and aggregate human capital across occupations are, respectively

$$\theta_{t+1} = \delta\theta_t^0 + (1-\delta)\theta_t\mu_t, \tag{18}$$

and

$$H_{t+1} = \delta H_t^0 + (1-\delta)H_t\mathcal{M}_t, \tag{19}$$

where in each period, the new workers' employment shares and human capital capital stocks are given by their optimal entry in the labor markets:

$$\theta_t^{0,j} = \frac{\left[\lambda_j \tau^{0,j} \left(-v_t^j\right)^{\frac{1}{1-\gamma}}\right]^{\alpha}}{\sum_{k=1}^{J} \left[\lambda_k \tau^{0,k} \left(-v_t^k\right)^{\frac{1}{1-\gamma}}\right]^{\alpha}},$$

and

$$H_t^{0,j} = \Gamma\left(1 - \frac{1}{\alpha}\right) \tau^{0,j}\, \lambda_j \left[\theta_t^{0,j}\right]^{1-\frac{1}{\alpha}},$$

where the vector τ^0 denotes the transferability of the a new worker's human capital into occupation j at the time of entering labor markets for the first time.

4.2.2 Firms' Optimization and Labor Demand

In this setting, productivity differences and the linearity of $q_t(x)$ ensures that, except for a set with measure zero, each of the tasks within an occupation will be provided either by only labor or by only machines, according to their comparative advantage. To see this, let w_t^j be the unitary price of effective labor j and r_t^M be the rental rate of a machine at time t. Because of perfect substitution, the minimum cost of producing one unit of task x in occupation j is

$$c_t^j(x) = \min\left\{\frac{w_t^j}{z_t^{H,j}(x)}, \frac{r_t^M}{z_t^{M,j}(x)}\right\}. \tag{20}$$

Clearly, the ratios between factor prices and productivities determine whether workers or machines will take care of a particular task in a particular occupation.[21] The minimized unitary cost of producing the tasks from every occupation j, C_t^j, is the solution of the program:

$$C_t^j = \min_{q_t^j(x)} \int_0^1 c_t^j(x)\, q_t^j(x)\, dx \quad \text{s.t.} \quad \left(\int_0^1 \left[q_t^j(x)\right]^{\frac{\eta-1}{\eta}} dx\right)^{\frac{\eta}{\eta-1}} = 1. \tag{21}$$

Having solved for C_t^j, the next step for the firm is to solve for C_t^Q, the minimized cost of producing the aggregate bundles of tasks, by optimally using the inputs from all occupations. That problem is simply

$$C_t^Q = \min_{\{Q_j\}_{j=1}^J} \left[\sum_{j=1}^{J} C_t^j Q^j\right] \quad \text{s.t.} \quad \left(\sum_{j=1}^{J} \left(Q^j\right)^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}} = 1. \tag{22}$$

Finally, given the rental price r_t^K for physical capital K_t, and the unitary cost of tasks C_t^Q, the

[21] Our task and production assignment is connects with Acemoglu and Restrepo (2018). In our setup the assignments of workers and machines to tasks is determined by factor prices and their relative productivities. As we show next, using properties of the Frechet distribution, we get very tractable and intuitive expressions.

competitive price of final goods is simply its minimized unitary production cost, i.e.:

$$P_t = \min_{K,Q} \left[r_t^K K + C_t^Q Q \right] \quad \text{s.t.} \;\; (K_t)^{\varphi} (Q_t)^{1-\varphi} = 1. \tag{23}$$

The following proposition characterizes the solution of all these optimization problems:

Proposition 2 *For all $j = 1, \ldots J$, assume that $z_t^{H,j}(x)$ are all distributed i.i.d. Frechet, with shape and scale parameters $\left\{\nu_j, A_t^{H,j}\right\}_{j=1}^J$. Similarly, for machines, assume that the productivities $z_t^{M,j}(x)$ are also distributed Frechet with shape and scale parameters $\left\{\nu_j, A_t^{M,j}\right\}_{j=1}^J$ across occupations j. Then:* ***(a)*** *The probability that at time t labor implements any task x in occupation j is*

$$\pi_t^{H,j} = \frac{\left(A_t^{H,j}/w_t^j\right)^{\nu}}{\left(A_t^{M,j}/r_t^M\right)^{\nu} + \left(A_t^{H,j}/w_t^j\right)^{\nu}}, \tag{24}$$

while the probability that the task is implemented by machines is

$$\pi_t^{M,j} = \frac{\left(A_t^{M,j}/r_t^M\right)^{\nu_j}}{\left(A_t^{M,j}/r_t^M\right)^{\nu_j} + \left(A_t^{H,j}/w_t^j\right)^{\nu_j}}. \tag{25}$$

(b) *The minimized unitary costs C_t^j of producing the bundle of tasks from each occupation j are*

$$C_t^j = \Gamma\left(1 + \frac{1-\eta}{\nu_j}\right)^{\frac{1}{1-\eta}} \left[\left(A_t^M/r_t^{M,j}\right)^{\nu_j} + \left(A_t^{H,j}/w_t^j\right)^{\nu_j}\right]^{\frac{-1}{\nu_j}}. \tag{26}$$

(c) *The minimized unitary costs C_t^Q of producing the aggregate bundle of tasks is*

$$C_t^Q = \left\{\sum_{j=1}^J \Gamma\left(1 + \frac{1-\eta}{\nu_j}\right)^{\frac{1-\rho}{1-\eta}} \left[\left(A_t^{M,j}/r_t^M\right)^{\nu_j} + \left(A_t^{H,j}/w_t^j\right)^{\nu_j}\right]^{\frac{\rho-1}{\nu_j}}\right\}^{\frac{1}{1-\rho}}. \tag{27}$$

(d) *The price of the final goods is*

$$P_t = \left[\varphi^{-\varphi}(1-\varphi)^{\varphi-1}\right] \left(r_t^K\right)^{\varphi} \left(C_t^Q\right)^{1-\varphi}. \tag{28}$$

4.2.3 Capital Owners

Given an initial level structures, $K_0 > 0$, and machines $M_0^j > 0$, for $j = 1, \ldots, J$, an initial financial position B_0 and a sequence of good prices, capital rental rates, and interest rates, $\left\{P_t, r_t^K, r_t^M, R_t\right\}_{t=0}^{\infty}$, the budget constraint of capital owners, at any period t, is

$$\frac{r_t^M}{P_t}\sum_{j=1}^J M_t^j + \frac{r_t^K}{P_t} K_t + \frac{R_t}{P_t} B_t = c_t^K + I_t^K + \sum_{j=1}^J I_t^{M,j} + B_{t+1}, \tag{29}$$

where the laws of motion for K_t and M_t^j are given by (13) and (14), respectively.

Under the conditions just stated, the program of consumption, investments and capital stocks, $\left\{c_t^K, I_t^K, I_t^{M,j}, K_{t+1}, M_{t+1}^j, B_{t+1}\right\}_{t=0}^{\infty}$, that maximizes (12) is characterized by a standard transversality condition, and three Euler equations that can be written as:

$$\frac{R_{t+1}}{P_{t+1}} = \beta^{-1}\left(\frac{c_{t+1}^K}{c_t^K}\right)^{\gamma}, \tag{30}$$

$$\frac{r_{t+1}^K}{P_{t+1}} = \frac{\frac{R_{t+1}}{P_{t+1}} - \left(1-\delta^K\right)}{\xi_t^K}, \tag{31}$$

$$\frac{r_{t+1}^M}{P_{t+1}} = \frac{\frac{R_{t+1}}{P_{t+1}} - \left(1-\delta^M\right)}{\xi_t^M}, \tag{32}$$

where (32) applies for all the machines in occupations $j = 1, \ldots J$.

4.2.4 Competitive Equilibrium

Having characterized the individual optimality conditions of all agents in the economy, we now define and characterize the competitive equilibria in this economy.

Definition 1 (Equilibrium) *Given an initial population of workers and their human capital,* $\left\{\theta_0^j,\, H_0^j\right\}_{j=1}^J$*, initial stocks of machines and other physical capital* $\left\{M_0^j\right\}_{j=1}^J$*,* K_0*, and an exogenous productivities sequence* $\left\{\left\{A_t^{M,j}, A_t^{H,j}\right\}_{j=1}^J\right\}_{t=0}^{\infty}$ ***an equilibrium*** *is **(i)** a* price system $\left\{w_t^j, P_t, r_t^K, r_t^M, R_t\right\}_{t=0}^{\infty}$*, **(ii)** individual worker occupation decisions* $\left\{v_t^j, \mu_t\right\}_{t=0}^{\infty}$*, **(iii)** individual firm tasks-allocation choices* $\left\{\pi_t^{M,j}, \pi_t^{H,j}\right\}_{t=0}^{\infty}$*, **(iv)** aggregate* vectors of workers and human capital across occupations, stocks of machines and other physical capital, $\left\{\theta_t, H_t, M_t^j, K_t\right\}_{t=0}^{\infty}$*, and, **(v)** aggregate output,* worker and human capital reallocations, and flows of investments and of consumption of the owners of capital, $\left\{Y_t,\, \mu_t,\, \mathcal{M}_t, I_t^K, I_t^M, c_t^K\right\}_{t=0}^{\infty}$ *such that: **(a)** Given* $\left\{w_t^j, P_t, r_t^K, r_t^M\right\}_{t=0}^{\infty}$*, the workers lifetime optimization* $\left\{v_t^j, \mu_t\right\}_{t=0}^{\infty}$ *are given by (15) and (16); the firms optimize production, i.e.* $\left\{\pi_t^{H,j}, \pi_t^{M,j}, P_t\right\}_{t=0}^{\infty}$ *are given by (24), (25), and (28); and capital owners invest optimally, i.e. according to (29), (31), and (32.) **(b)** factor markets clear in every period t, and **(c)** the population of workers and human capital allocation evolve according to (18) and (19.)*

We now characterize the aggregate output and the prices of human and physical capital that arise from the market-clearing conditions in every period given an exogenous level for productivities $\left\{A_t^j, A_t^M\right\}$, and pre-determined levels of aggregate supplies H_t^j, M_t and K_t.

4.3 Static Market Clearing Conditions

The following proposition characterizes the intratemporal equilibrium conditions determining prices and production, taking as given the stocks of physical and human capital.

Proposition 3 ***Aggregation, Intratemporal Equilibrium.*** *Given pre-determined aggregate variables,* $\left\{K_t,\, M_t^j,\, H_t^j\right\}$*, the intratemporal competitive equilibrium condition imply that:*

(a) *the output of tasks of all occupations* j*,* Q_t^j*, the aggregate bundle of tasks,* Q_t*, and the aggregate output of goods,* Y_t*, are respectively given by:*

$$Q_t^j = \Gamma\left(1+\frac{1-\eta}{\nu_j}\right)^{\frac{1}{\eta-1}}\left[\left(A_t^{M,j}M_t^j\right)^{\frac{\nu_j}{1+\nu_j}}+\left(A_t^{H,j}H_t^j\right)^{\frac{\nu_j}{1+\nu_j}}\right]^{\frac{1+\nu_j}{\nu_j}}, \tag{33}$$

$$Q_t = \left[\sum_{j=1}^{J}\Gamma\left(1+\frac{1-\eta}{\nu_j}\right)^{\frac{\rho-1}{\rho(\eta-1)}}\left[\left(A_t^{M,j}M_t^j\right)^{\frac{\nu_j}{1+\nu_j}}+\left(A_t^{jH}H_t^j\right)^{\frac{\nu_j}{1+\nu_j}}\right]^{\frac{(1+\nu_j)(\rho-1)}{\nu_j\,\rho}}\right]^{\frac{\rho}{\rho-1}}, \tag{34}$$

$$Y_t = (K_t)^{\varphi}\left[\sum_{j=1}^{J}\Gamma\left(1+\frac{1-\eta}{\nu_j}\right)^{\frac{\rho-1}{\rho(\eta-1)}}\left[\left(A_t^{jM}M_t^j\right)^{\frac{\nu_j}{1+\nu_j}}+\left(A_t^{jH}H_t^j\right)^{\frac{\nu_j}{1+\nu_j}}\right]^{\frac{(1+\nu_j)(\rho-1)}{\nu_j\,\rho}}\right]^{\frac{(1-\varphi)\rho}{\rho-1}} \tag{35}$$

(b) *The equilibrium real rental rates of structures,* $\rho_t^K \equiv r_t^K/P_t$*, and of machines in all occupations* j $\rho_t^{M,j} \equiv r_t^{M,j}/P_t$ *and the real unitary wages of workers* $\omega_t^j \equiv w_t^j/P_t$*, are, respectively*

$$\rho_t^K = \varphi\frac{Y_t}{K_t}, \tag{36}$$

$$\rho^{M,j} = (1-\varphi)\,\Gamma\left(1+\frac{1-\eta}{\nu_j}\right)^{\frac{1}{\eta-1}}\left(\frac{A_t^{M,j}M_t^j}{Q_t^j}\right)^{\frac{\nu_j}{1+\nu_j}}\left(\frac{Q_t^j}{Q_t}\right)^{\frac{\rho-1}{\rho}-\frac{\nu_j}{1+\nu_j}}\frac{Y_t}{M_t^j}, \tag{37}$$

$$\omega^{j} = (1-\varphi)\,\Gamma\left(1+\frac{1-\eta}{\nu_j}\right)^{\frac{1}{\eta-1}}\left(\frac{A_t^{M,j}H_t^j}{Q_t^j}\right)^{\frac{\nu_j}{1+\nu_j}}\left(\frac{Q_t^j}{Q_t}\right)^{\frac{\rho-1}{\rho}-\frac{\nu_j}{1+\nu_j}}\frac{Y_t}{H_t^j}. \tag{38}$$

These simple aggregation results, can directly be used to solved for the income share of the different factors of production. In particular, the total capital (structures+machines) output share is given by

$$\frac{\rho_t^K K_t+\sum_{j=1}^{J}\rho_t^{M,j}M_t^j}{Y_t} = \varphi+(1-\varphi)\sum_{j=1}^{J}\Gamma\left(1+\frac{1-\eta}{\nu_j}\right)^{\frac{1}{\eta-1}}\left(\frac{A_t^{M,j}M_t^j}{Q_t^j}\right)^{\frac{\nu_j}{1+\nu_j}}\left(\frac{Q_t^j}{Q_t}\right)^{\frac{\rho-1}{\rho}-\frac{\nu_j}{1+\nu_j}},$$

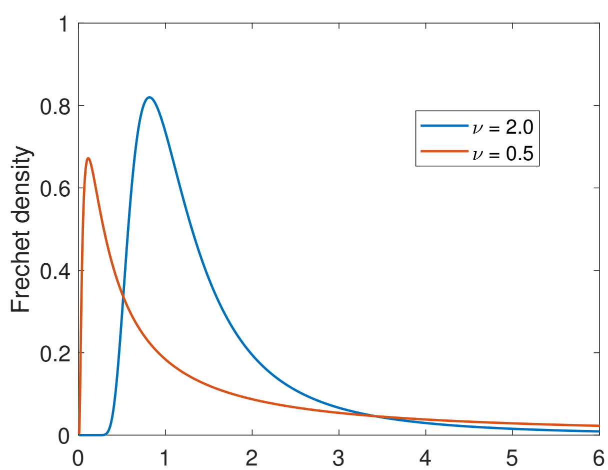
which endogenously responds to the relative supply of human and physical capital and their relative productivity levels. To illustrate this point more clearly, consider the special case in which Q_t is a Cobb-Douglass, i.e., $\rho = 1$, over all occupations outputs Q_t^j with shares ϱ_j. In such a case, the capital income share boils down to

$$\frac{\rho_t^K K_t+\sum_{j=1}^{J}\rho_t^{M,j}M_t^j}{K_t} = \varphi+(1-\varphi)\sum_{j=1}^{J}\varrho_j\frac{\left(A_t^{M,j}/r_t^M\right)^{\nu_j}}{\left(A_t^{M,j}/r_t^M\right)^{\nu_j}+\left(A_t^{H,j}/w_t^j\right)^{\nu_j}}, \tag{39}$$

where we have used the expressions in Proposition 2 that characterizes the output share of machines in each of the occupations j. This expression highlights how levels of technology, wages and rental rates endogenously drive the capital output share. Extending Acemoglu and Restrepo (2018) to a tractable setting with multiple types of occupations and labor, the labor share of our economy depends on how efficient are machines in performing different tasks relative to labor. For example, an increase in the productivity of machines, $A_t^{M,j}$, across at least some of the occupations, or a reduction in the rental rates of machines, r_t^M, would lead to an overall decrease in the labor share.

An additional important aspect of Proposition 3 is that we can characterize the aggregate production of our economy. It is worth highlighting that different factors of production differ in their comparative advantage in the production of individual tasks, yet we can aggregate all the task that are completed by the different factors in a tractable way. Our task-based approach connects closely with a large macroeconomic literature on skill-biased technical change which directly assumes an aggregate CES production function using different factors (see for example Krusell et al. (2000)). In this way, our proposed model of production provides a micro-foundation for the CES aggregate production functions typically used in this literature. Note that the parameters ν_j, which affect the distribution of the productivity of machines and labor in the production of different tasks, are directly linked to the aggregate elasticity of substitution between capital and labor in the production of the output of an occupation, which is $1 + \nu_j$. The intuition for this can be easily grasped by looking at Figure 6, which shows two different distribution for task productivity $z^{H,j}$, one with a relatively large value of $\nu_j = 2$ and one with a lower value of $\nu_j = 0.5$, and equal scale parameters. Note that for the distribution with a lower ν_j, the dispersion in the productivity of different factors is larger (fatter tail), while is smaller in the other case. In other words, factors tend to have a more similar productivity –similar comparative advantage– with larger ν_j. This implies that, for a similar change in factor prices, there will be more substitution between labor and machines in the case of larger ν_j. Thus, in the aggregate, the elasticity of substitution between factors is governed by this parameter of the shape of the Frechet distribution of productivities.

Figure 6: Different dispersion in comparative advantages of labor and machines



4.4 Dynamics

We now consider the dynamic behavior of the economy. We first consider the time-invariant equilibria, when the economy follows a balanced-growth paths (BGP). We then consider the behavior of the economy outside the BGP, that is, the dynamic equilibrium responses of the economy to changes in, for example, the underlying productivities of both labor and machines.

4.4.1 Balanced Growth Paths (BGP)

Consider now the long-run behavior of an economy in which the productivity of different types of labor grows at a common and exogenous gross rate G_A over time, while that of machines stays constant. That is, $A_{t+1}^{Hj} = G_A \, A_t^{H,j}$ and $A_{t+1}^{Mj} = A_t^{Mj}$ in the long-run. An equilibrium would accrue where all the real unit wages grow at the same rate G_A but real rental rates are constant. In this equilibrium, aggregate human capital will be constant, but physical capital (machines and structures), labor earnings and total output grow at rate G_A.

Assume that the productivity of all types of labor grow at a constant rate $G_A \geq 1$, i.e.:

$$A_t^{Mj} = a^{Mj}, \text{ and, } A_t^{Hj} = a^{Hj} \, (G_A)^t,$$

for all $j = 1, ..., J$ and $t \geq 0$, for some exogenous constants $G_A > 0$ and a^{Hj}, a^{Mj}. In addition to the rate G_A, the earnings of individual workers along a BGP would grow, on average, at the endogenously determined rate G_H. In what follows, we restrict attention to the case when $G_H < (1-\delta)^{-1}$, so that the total weight of earnings of cohorts fall as they age.[22] Hence, the aggregate human capital of the country settles to a time-invariant level, $H = \delta H^0 \left[I - (1-\delta)\mathcal{M}\right]^{-1}$, and the rate of growth of the economy is the exogenous rate G_A.

The dynamic programming results in the previous section easily extend to the case where real wages grow at a common constant rate G_A. In this case, lifetime utilities per unit of human capital grow at rate $G_A^{1-\gamma}$. This common trend may impact occupation choices by changing the effective discount factor of the workers. Define $v_t^j = \bar{v}^j \, (G_A^{1-\gamma})^t$ and $\omega_t^j = \bar{\omega}^j \, (G_A)^t$, where $\bar{v}$ and $\bar{\omega}$ are constant along a BGP. The problem of the worker in a BGP for $\gamma > 1$ takes the form

$$\bar{v}^j \, (G_A^{1-\gamma})^t = \frac{(\bar{\omega}^j \, (G_A)^t)^{1-\gamma}}{1-\gamma} - \beta \, \Gamma \left(1 - \frac{1-\gamma}{\alpha}\right) \left[\sum_{\ell=1}^{J} (\chi_{j,\ell} \, \bar{v}^j \, (G_A^{1-\gamma})^{t+1})^{\frac{\alpha}{1-\gamma}} \, (\tau_{j\ell}\lambda_\ell)^\alpha\right]^{\frac{1-\gamma}{\alpha}},$$

which can be further simplified to,

$$\bar{v}^j = \frac{(\bar{\omega}^j)^{1-\gamma}}{1-\gamma} - \beta (G_A^{1-\gamma}) \, \Gamma \left(1 - \frac{1-\gamma}{\alpha}\right) \left[\sum_{\ell=1}^{J} (\chi_{j,\ell} \, \bar{v}^j)^{\frac{\alpha}{1-\gamma}} \, (\tau_{j\ell}\lambda_\ell)^\alpha\right]^{\frac{1-\gamma}{\alpha}}. \tag{40}$$

[22] Economies in which $G_H(1-\delta) > 1$ have the very unappealing feature that asymptotically, the aggregate labor income would be concentrated in a measure zero of workers with infinite age.

This is an expression very similar to that of Theorem 1.[23] The formulae for μ and $\mathcal{M}$ are essentially those of Section 2, but using $\bar{v}$. The average growth rate of each individual worker's human capital over his life-cycle, G_H, is governed by the Perron root of $\mathcal{M}$, which, as explained there, is unique, real and strictly positive.

The Euler equations (31) and (32) of capital owners require that the rental rates of both forms of physical capital along a BGP satisfy

$$\rho^K = \frac{R/P - \left(1-\delta^K\right)}{\xi^K}, \tag{41}$$

$$\rho^M = \frac{R/P - \left(1-\delta^M\right)}{\xi^M}, \tag{42}$$

The analysis that follows applies equally to *small open economies*, where the real interest rate is exogenously given, i.e.: $R/P = R^*/P^*$, and to *closed economies*, where the interest rate in a BGP is also exogenously given by $R/P = \beta^{-1}\left(G_A\right)^{\gamma}$. In both cases, the rental rate of structures and machines is pinned down by the Euler equations (41) and (42.)

Before proceeding to a complete definition, we introduce some additional notation. Along a BGP, we have $I_t^K = i^K\left(G_A\right)^t$, $I_t^{M,j} = i^{M,j}\left(G_A\right)^t$, $K_t = k\left(G_A\right)^t$, $M_t^j = m^j\left(G_A\right)^t$, and $Y_t = y\left(G_A\right)^t$, where i^K, $i^{M,j}$, k, m^j, and y, are constants that depend on parameters and growth rates.

We use the following definition of a BGP:

Definition 2 *Given an exogenous productivity growth rate G_A, a* ***Balanced Growth Path*** *is a pair of constant rental rates, $\left(\rho^K, \rho^M\right)$, a constant real interest rate, R/P, a vector of constant detrended wages, $\bar{\omega}^j$, the average growth rate of the workers' human capital, G_H, a positive vector $h \in \mathbb{R}_{++}^J$ of shares of aggregate human capitals, and positive constants $\left(k, m^j, i^K, i^{Mj}, y\right)$ related to investment, capital and output, and individual solutions for the workers problems $\{\bar{v}, \mu, \mathcal{M}\}$ such that:* ***(a)*** *$\left(\rho^K, \rho^M, \bar{\omega}^j\right)$ solve the intratemporal conditions (36), (37) and (38) for h, k, m^j;* ***(b)*** *The growth rate G_H is the Perron root of $\mathcal{M}$ and h is proportional to $H = \delta H^0\left[I-(1-\delta)\mathcal{M}\right]^{-1}$.* ***(c)*** *The rental rates of structures and machines ρ^K, ρ^M satisfy the Euler equations (41) and (42).* ***(d)*** *Given $\bar{\omega}^j$, $\{\bar{v}, \mu\}$ solves the individual worker's optimal occupation choice problem and $\mathcal{M}$ is the associated transition function for aggregate human capital.*

Theorem 2 *Consider an economy with:* ***(a)*** *strictly positive vector of labor productivity levels, A^M and $\left\{A_t^j\right\}_{j=1}^J$, with gross growth $G_A \geq 1$;* ***(b)*** *rental rates of machines and structures ρ_{bgp}^M,ρ_{bgp}^K are that are either exogenously given or determined by G_A and the Euler equations of the capital owners;* ***(c)*** *the parameters are such that the worker's problem with growth, (40), is well defined. Then:* ***(1)*** *There exists a BGP in which a time invariant $\{\bar{v}, \mu\}$ solve the individual worker's problem and induce constant transition matrices $\mu, \mathcal{M}$ and an invariant distribution of workers attains, i.e., $\theta = \delta\theta^0\left[I-(1-\delta)\mu\right]^{-1}$. Moreover, (i) if the implied Perron root satisfies $G_H < (1-$*

[23] All the existence and uniqueness proofs of the value function extend to this case for variable $\bar{v}$, but adjusting the discount factor to be $\beta(G_A^{1-\gamma})$.

$\delta)^{-1}$*, then the aggregate human capital remains constant,* $H = \delta H^0 \left[I - (1-\delta)\mathcal{M}\right]^{-1}$*, and output, capital and wages,* $\left\{Y_t, K_t, M_t^j, w_t^j\right\}_{t=0}^{\infty}$*.* ***(i)*** *if instead, the Perron root satisfies* $G_H > (1-\delta)^{-1}$*, then the aggregate human capital remains grows at the rate* $(1-\delta)G_H$*, wages grow at the rate* G_A*, and output and capital grow at the rate* $G_A G_H(1-\delta)$*.* $\left\{Y_t, K_t, M_t^j, w_t^j\right\}_{t=0}^{\infty}$*.* ***(ii)*** *If there are unique ratios* H^k/M^j *associated to* ρ_{bgp}^M*,*ρ_{bgp}^K*, then the BGP is unique.*

Proving uniqueness of the BGP has been a more elusive task for the general case. However, we show in the appendix that it boils down to examining the solution to a system of J equations on J unknowns, which can be easily done computationally.[24]

4.5 Transitions: Dynamic Hat Algebra

Having established the conditions for a BGP, in this section we examine the implied dynamics of the model outside a BGP. To this end, in this section we extend the Dynamic Hat Algebra (DHA) methods of Caliendo, Dvorkin, and Parro (2019) to a model with general CRRA preferences, human capital accumulation and endogenous growth.

Proposition 4 ***Dynamic Hat Algebra****. If initial allocations of the new cohort of workers and human capital across occupations at period* $t=0$*,* $\theta_0^{0,j} > 0, H_0^{0,j} > 0$*, transition matrices of workers and human capital,* $\mu_{-1}, \mathcal{M}_{-1}$*, initial ratio of current vs permanent income* Φ_0^j*, lifetime and factor payments/shares,* $\pi_0^{H,j}, \pi_0^{M,j}$ *and the share of occupation* j *in non-structures value added,* ϱ_0^j*,* $\forall\, j$*, are all observed, and the values for the discount factor* β*, the CRRA coefficient* γ*, the depreciation rates* δ^M*,* δ^K*, the share of structures in aggregate income,* φ*, the elasticity of substitution between different occupations in production,* ρ*, and the curvature parameters* α *and* ν_j*, are known (estimated or calibrated), then:* ***(a)*** *the sequential equilibrium of this economy can be written in changes relative to the BGP.* ***(b)*** *Given an unanticipated change in machines or workers' productivity levels,* $\left(\left\{A^{H,j},\, A^{M,j}\right\}_{j=1}^{J}\right)$*, or price of equipment (investment efficiency),* ξ_t^M*, we can compute the sequential equilibrium of this economy in changes.* ***(c)*** *For both (a) and (b), it is not necessary to know the levels of a large number of parameters related to mobility frictions (*τ*,* ξ*), or productivities (*A^H*,* A^M*,* λ*,* ξ^K*,* ξ^M*).*

The proof of the proposition is in Appendix B, where we also lay down the equations that describe the model in changes relative to a BGP.

Using dynamic-hat-algebra methods is particularly convenient for the computation of the model for two reasons. First, the levels of a large set of parameters, like τ, χ, λ, or the initial levels of productivity, A^M, A^H, ξ^M ξ^k, are not needed to calibrate the model or to perform counterfactual analysis, only the *changes* in these variables, to the extent that they change, are required. This implies that the calibration exercise is less demanding. Second, the level of many of the model's

[24] We have an iterative algorithm to check uniqueness. In our computational exercises and under our preferred calibration, we use different initial guesses and always obtain the same BGP.

endogenous variables are not needed and the initial and terminal values for many endogenous variables expressed in *changes* are easy to characterize.

5 Quantitative Analysis

In this section, we use our model to quantitatively explore whether a sequence of task biased technological changes (tbtc) can jointly account for two salient trends in the U.S. economy, as observed during the last forty years. First, the output share of labor has declined substantially, around 5% as documented by Karabarbounis and Neiman (2013) since 1980. Second, there has been a remarkable polarization in employment and earnings in U.S. labor markets and an increase in overall wage inequality, as summarized by Acemoglu and Autor (2011). Like these authors, we highlight technological changes as an explanation for production shifts from workers to machines. However, our emphasis is on the ensuing reallocation of workers and human capital from some occupations that are losing their race with machines to other occupations.

We first describe our sources of data and our calibration strategy. In particular, we justify our use of observed data for the U.S. in the 1970s as an initial equilibrium BGP, explaining the moments we match. Then, we describe how we set the values for some key parameters. Next, we describe how we use data on the relative price of equipment and occupation shares to calibrate the sequence of task-biased innovations hitting the economy since the early 1980s. We ascertain how much the model replicates the increased inequality observed in the data since 1980 by comparing the results of the model with tbtc vs. the underlying economy without those shocks.

5.1 Data and Initial Equilibrium

We first describe our calibration strategy for the parameters related to the worker's problem. Then, we describe our calibration of the parameters linked to production.

We assume that the U.S. economy was in a BGP by the end of the 1970s and calibrate the initial BGP of our model to match the equilibrium conditions to those of the U.S. at that time. For this we need a reliable source of microdata on occupational mobility and earnings dynamics with a panel dimension. The panel dimension is needed to compute measures of occupational mobility of workers and also how workers' earnings change with occupational moves. Following Kambourov and Manovskii (2013), we use the Panel Study of Income Dynamics (PSID), which has a representative sample of the U.S. economy and is widely used in the labor and macroeconomics literature. A key advantage of the PSID is that information on occupations between 1968 and 1980 has been corrected for occupation miss-classifications that lead to spurious occupational switches. As discussed in Kambourov and Manovskii (2013), occupation miss-classification can be an important source of biases for measures of occupational mobility.

Typically, the choice of the level of disaggregation for occupations has to balance computational costs and sample sizes of available data. The latter is the limiting constraint for our exercise because of the relatively small sample of the PSID.[25] We calibrate our model to nine broad occupations: (1) Management, business, and financial operations occupations; (2) Professional and related occupations; (3) Service occupations; (4) Sales and related occupations; (5) Office and administrative support occupations; (6) Construction and extraction occupations; (7) Installation, maintenance, and repair occupations; (8) Production occupations; and (9) Transportation and material moving occupations.[26] For brevity, in what follows, we refer to these occupations simply as Managers, Professionals, Service, Sales, Office, Construction, Repair, Production, and Transportation, respectively.

We calibrate some of the parameters in our model exogenously to standard values in the literature. We take a period in our model to represent a year and workers discount the future at rate $\beta = 0.95$, a standard value for an annual model. We set the risk aversion parameter to 2, a usual value in the macroeconomics literature. Workers have an stochastic lifetime with a probability of exiting the labor market (due to death, disability, non-participation decisions) of 3% per year. Then, on average, individuals in the model work for an average of 33 years.

The computational strategy using the dynamic hat algebra of Proposition 4 requires that we map certain endogenous variables of the model in the initial period directly to moments in the data. We now describe in detail how we proceed.

In our model, workers start their lifetime and enter the labor market with a level of human capital and earnings that is on average lower than that of older workers. We map the average earnings and occupational employment of the entering cohort (variables θ_0^0 and H_0^0) to those of workers with ages between 22 and 30 years in the PSID data for all years before 1980. Table 2 shows the distribution of the entering cohort over occupations and their human capital. Relative to older workers, a much lower fraction of the new workers enters in managerial occupations (7% vs. 14%) and higher fractions enter in production, office and construction. The implied human capital levels H_0^0 are also remarkably lower, as we discuss below.

Table 2: U.S. in the 1970s: BGP Shares of Young Workers and Human Capital Across Occupations

	entering cohort								
	Managers	Profess.	Service	Sales	Office	Constr.	Repair	Production	Transport
Workers (θ_0^0):	0.07	0.13	0.12	0.05	0.15	0.08	0.05	0.21	0.13
Human Capital (H_0^0):	0.08	0.14	0.08	0.05	0.14	0.09	0.06	0.22	0.13

Along the BGP, workers' earnings grow at rate $G_A \times G_H$ in our model. Then, using data only on the growth rate of earnings, we cannot identify these two components separately. However, we assume that the human capital of the entering cohort in the BGP does not grow, thus the only

[25] Our model is very tractable and the computational demands at this level of disaggregation are very low. While we would like to substantially increase the numbers of occupations we use, the lack of suitable data prevents us from doing so.

[26] We exclude farming, fishing, and forestry occupations, because they account for a minimal share of U.S. employment and the PSID includes very few observations in the sample. We also exclude all military occupations.

source of earnings growth for the entering cohorts is the growth rate of wages, i.e. G_A. We use the growth rate of earnings of workers between 22 and 30 to calibrate $G_A = 1\%$.[27]

Then, using the PSID data for all workers between 22 and 65 years old, we estimate the yearly occupational mobility matrix (variable μ_{-1} in the model) using the Poisson Maximum-Likelihood methods proposed by Silva and Tenreyro (2006).[28]

Table 3 reports the estimated matrix μ for the U.S. economy in the 1970s. As expected, there is substantial persistence in occupation choices, as can be seen in the diagonal of the estimated matrix μ which exceedingly dominates the off-diagonal terms. There is also substantial variation in the degree of persistence across the different occupations. On the upper end, 88.4% of production workers remain in those occupations for the following year. On the lower end, just 76.5% of sales workers stay in those jobs from one year to the next.

Table 3: U.S. in the 1970s: Workers Occupational Transition Matrix μ_{-1}

From \To	Managers	Profess.	Service	Sales	Office	Constr.	Repair	Production	Transport
Managers	0.88	0.02	0.02	0.01	0.02	0.01	0.01	0.02	0.02
Profess.	0.02	0.86	0.02	0.01	0.02	0.01	0.01	0.02	0.02
Service	0.02	0.02	0.88	0.01	0.02	0.01	0.01	0.02	0.02
Sales	0.03	0.03	0.03	0.76	0.03	0.02	0.02	0.04	0.03
Office	0.02	0.02	0.02	0.01	0.85	0.01	0.01	0.02	0.02
Constr.	0.03	0.03	0.03	0.01	0.02	0.79	0.02	0.03	0.03
Repair	0.03	0.03	0.03	0.01	0.02	0.02	0.81	0.03	0.03
Production	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.88	0.02
Transport	0.02	0.02	0.02	0.01	0.02	0.01	0.01	0.02	0.87

Similarly, we estimate the matrix $\mathcal{M}$, using the information on earnings dynamics for occupational switchers and stayers. Consistent with the BGP in our model and the previous discussion on the calibration of parameter G_A, unitary wages are all growing at rate G_A in the BGP. As is usually the case with Roy models, we cannot distinguish the (relative or detrended) level of unit wages and the total number of efficiency units of labor (or units of human capital) across occupations since observed earnings is the product of both. We proceed by normalizing the vector of detrended unitary wages to be all equal to one and estimate the matrix $\mathcal{M}_{-1}$ by the product of the matrix of average earnings changes for occupational switchers and stayers by occupation and the matrix μ_{-1} discussed before, as implied by our model.[29] The estimation of the matrix of average earnings changes uses the same data and estimation procedure used in the estimation

[27]For this we use data from the CPS since it has a much larger sample size and we do not need the panel dimension or information about occupations. Since the economy faces several deep recessions in the late 70s and early 80s, we use the average growth rate in earnings for the years 1983 to 1989.

[28]In this way, the small size of our sample for some transitions would have less influence in our estimated moments than on the estimates directly using shares from the data, i.e., using bin estimators. We obtain similar results using a logit estimator as in Kambourov and Manovskii (2008).

[29] This normalization on the vector of (relative or detrended) wages is inconsequential for the results since any other normalization for wages would lead to a different level of the individual components of the vector of human capital H.

of the matrix of occupational switching μ_{-1}. To remove the effect of wage growth from G_A we divided the matrix of average earnings changes by it. Table 2 reports the matrix $\mathcal{M}_{-1}$ estimated from the data in the initial BGP.

Table 4: U.S. in the 1970s: Human Capital Occupational Transition Matrix $\mathcal{M}_{-1}$

From \To	Managers	Profess.	Service	Sales	Office	Constr.	Repair	Production	Transport
Managers	0.89	0.02	0.02	0.01	0.01	0.01	0.01	0.02	0.02
Profess.	0.02	0.89	0.02	0.01	0.02	0.01	0.01	0.02	0.02
Service	0.02	0.02	0.90	0.01	0.02	0.01	0.01	0.02	0.02
Sales	0.04	0.03	0.04	0.78	0.03	0.02	0.02	0.04	0.03
Office	0.03	0.02	0.02	0.01	0.87	0.01	0.02	0.03	0.02
Constr.	0.03	0.03	0.03	0.01	0.02	0.81	0.02	0.03	0.03
Repair	0.03	0.02	0.03	0.01	0.02	0.02	0.82	0.03	0.03
Production	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.89	0.02
Transport	0.02	0.02	0.02	0.01	0.02	0.01	0.01	0.02	0.88

As expected, there are similarities between the matrices $\mathcal{M}_{-1}$ and the matrix μ_{-1}. However, recall that $\mathcal{M}_{-1}$ is not a stochastic matrix. Moreover, notice that the ratio between each entry of the matrix $\mathcal{M}_{-1}$ with the corresponding entry of the matrix μ_{-1} gives an estimate of the expected (average) evolution of human capital for the occupational switchers, conditional on switching. The range for this ratio is between 0.91 and 1.14.

From the estimated transition matrices μ_{-1} and $\mathcal{M}_{-1}$, together with the distribution of employment and human capital of young workers (entering cohort) over occupations and the expressions in Proposition 1, we can compute the BGP shares of workers, θ_0, and of aggregate human capital, H_0, distributed across occupations. Note that, since earnings grow over the lifetime of individuals, the level of human capital of all workers is, on average, higher than that of the entering cohort. Table 5 shows these estimates.

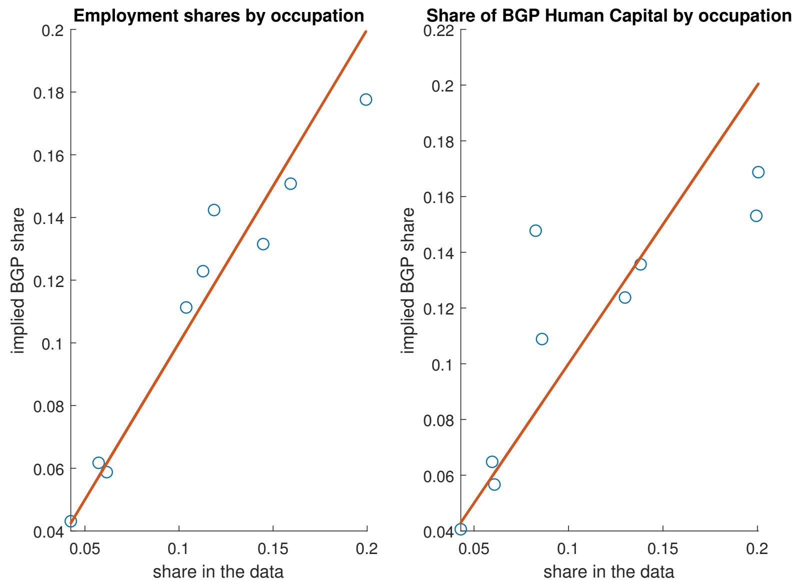
Table 5: U.S. in the 1970s: BGP Shares of Workers and Human Capital Across Occupations

	Managers	Profess.	Service	Sales	Office	Constr.	Repair	Production	Transport
Workers (θ_0):	0.14	0.12	0.15	0.04	0.11	0.06	0.06	0.18	0.13
Human Capital (H):	0.42	0.37	0.41	0.11	0.30	0.16	0.18	0.47	0.34

We can informally test the assumption that the economy is initially in a BGP by comparing the actual data on employment shares, θ_0, and earnings share –or share of human capital by occupation– with those implied by the model using the expressions in Proposition 1 together with the estimated transition matrices and the moments for the entering cohort. Figure 7 provides the two comparisons. We can see that the allocations in the data and the ones implied by the mobility matrices are not only very highly correlated but also of roughly the same magnitude as they lay very close to the 45-degree line.

The parameter α directly affects the dynamics of earnings and, other things equal, has a direct incidence on the amount of earnings inequality in the cross-section of all workers in the initial

Figure 7: Data vs BGP implied allocations



BGP. We assume a value of 13 which, for values of λ close to one, implies that permanent earnings shocks at the individual level do not have a large variance, consistent with the empirical literature on earnings dynamics (Lillard & Willis, 1978; MaCurdy, 1982).[30] In particular, in our model, the ϵ shocks will have a permanent effect on earnings. The variance of this permanent income shock given a value of $\alpha = 13$ is 0.01, which is very close to the estimates in Heathcote, Perri, and Violante (2010) for the periods before 1980.[31] Over time, however, the accumulation of these shocks can generate substantial inequality in the cross-section. In the next subsections we discuss how much initial inequality we get given our our calibration for α and other parameters.

Finally, as we discuss in Appendix B, our dynamic hat algebra method requires that we calibrate the ratio of current vs permanent income, Φ_0^j for the initial BGP. Given our normalization for wages in the initial BGP, we can compute values for Φ_0^j as functions of μ_{-1} and $\mathcal{M}_{-1}$, together with parameters β, α, and γ.

We now describe the calibration of parameters related to investment and production. To simplify the analysis, we take the (gross) real interest rate, R_t/P_t, to be exogenous (small open economy assumption) at the value of $(\beta^{-1} \times G_A^{1-\gamma}) = 1.05\%$ per year.[32] We set the depreciation rates of structures and investment to be $\delta^K = 0.05$ and $\delta^M = 0.125$, which are the values used by Greenwood et al. (1997); Krusell et al. (2000). Moreover, we follow those authors and calibrate the

[30]Note that, this large literature on random income processes abstracts from selection of workers over occupations due to comparative advantage, and from selection over most favourable shocks. We do account for this in our work.

[31]See Figure 18 in Heathcote et al. (2010), where the estimate of the variance of permanent income shocks, in levels, hoovers around 0.01.

[32]The small open economy assumption implies that investment of structures and machines will be perfectly elastic. Note however, that the rental rate of machines will respond to changes in the price of equipment, ξ_t^M, changing the total number of machines used in production. In the next subsection we discuss the exogenous changes in the price of equipment we use in our model.

share of structures in production, φ, to be 0.13. Last, we calibrate the elasticity of substitution of different occupations in production, ρ to be one, which is consistent with estimates in Goos et al. (2014).[33]

Similar to the problem of the worker, our quantitative strategy requires that we obtain values for the shares $\pi^{M,j}, \pi^{H,j}$ of machines and labor in the value added of each occupation, and also the share of each occupation in the non-structure value added, ϱ^j. We can use the equilibrium conditions of our model to link these endogenous variables with the data. First, note that in our model the total payments to labor in occupation j in the initial BGP relative to all income is, $\frac{w_0^j\, H_0^j}{P_0\, Y_0} = (1-\varphi)\, \varrho_0^j\, \pi_0^{H,j}$. Consider the following decomposition,

$$\frac{w_0^j\, H_0^j}{P_0\, Y_0} = \left(\frac{\sum_{k=1}^{J}(w_0^k\, H_0^k)}{P_0\, Y_0}\right) \frac{w_0^j\, H_0^j}{\sum_{k=1}^{J} w_0^k\, H_0^k}.$$

The first term on the right equals all the payments to labor relative to all income in the economy, i.e. the labor share. The last term is the share of occupation j earnings in all earnings. In the data, the labor share in the 1970's is around 0.66 and our calibrated value for φ is 0.13. Thus, we can use the share of occupation j in total earnings from the data to get,

$$\varrho_0^j\, \pi_0^{H,j} = \left(\frac{0.66}{0.87}\right) \frac{w_0^j\, H_0^j}{\sum_{k=1}^{J} w_0^k\, H_0^k}.$$

Unfortunately, we cannot identify separately these two variables using only this moment. Since both of them are a share, each has to be between zero and one, and the sum of ϱ^j over occupations has to be one. We make the assumption that in the initial BGP, $\pi_0^{H,j} = 0.66/0.87$ for all occupations.[34] Therefore, under this assumption, the share of earnings of occupation j directly identifies the share of occupation j value added in total production.

Table 6 contains all the information discussed in this section about the calibration of the parameters of the model or the value of some variables in the initial equilibrium. The last piece of information not discussed so far is the calibration of the parameters ν^j which, in the aggregate, directly affect the elasticity of substitution between labor and machines in production. The next section discusses the calibration of these parameters.

[33] While Goos et al. (2014) model is different from ours, we performed some robustness checks for alternative values of this parameter. Critically, the value of ρ is not overly important, but it must be that for a given change in factor prices is easier to substitute factors of production than to substitute occupations. Our parameter values are always on this region.

[34] While the share of labor and machines in each occupation may be different in the data, the incidence of machines and automation technology in 1970s was arguably low. If so, it would hardly explain variations in the output shares across the different occupations, which would support our assumption.

Table 6: Summary of calibration strategy & parameters

parameter / variable		value
time period: 1 year		
No. of occupations	J	9
yearly model, discount factor	β	0.95
labor market exit prob.		0.03
risk aversion	γ	2
shape Frechet workers	α	13
exogenous growth rate of labor prod.	$A_t^{H,j}$	1.01
employment share of new cohort	θ_0^0	see Table 2
human capital of new cohort	H_0^0	see Table 2
occupational mobility matrix	μ_{-1}	see Table 3
human capital transition matrix	$\mathcal{M}_{-1}$	see Table 4
employment share of all workers	θ_0	see Table 5
human capital of all workers	H_0	see Table 5
ratio of current vs permanent income	Φ_0^j	see Appendix B
depreciation rate machines	δ^M	0.125
depreciation rate structures	δ^K	0.05
share of structures	φ	0.13
elasticity between occupations	ρ	1.00
small open economy - real rate	R_t/P_t	$\beta^{-1}(G_A)^\gamma$
share of occ j value added	ϱ_0^j	$(H_0^j \,/\, \sum_k H_0^k)$
share of labor in occ j value added	$\pi_0^{H,j}$	0.76 $\forall\, j$

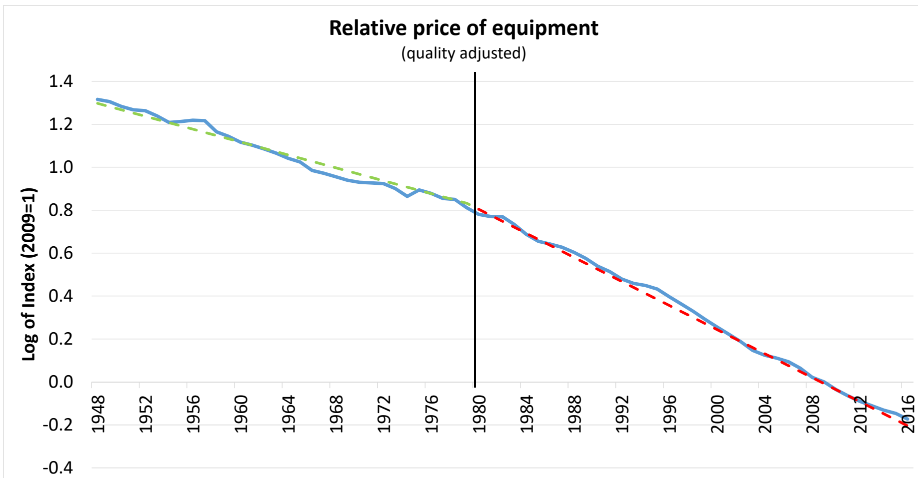
5.2 Task-biased Technological Shocks

Having set up the economy in a BGP as of 1980, we now use the model to capture the response to a long-lasting episode of task biased (tbtc) that may be asymmetric across the workers in different occupations. To this end, we use the information on the relative price of equipment to calibrate the changes in ξ_t^M. Recall that the Euler equation for the investment in equipment, (32), implies an inverse relationship between the rental rate of machines r_t^M and investment-specific productivity ξ_t^M. Thus, we follow the literature on skill-biased technical change (Greenwood et al., 1997; Krusell et al., 2000) and use data on the price of the investment in equipment relative to the price of consumption to calibrate the evolution of task biased technology from 1980 to 2010. This is the unanticipated tbtc shock we feed to our economy. We assume perfect foresight for individuals in the economy. Specifically, we start the economy in 1980, our time $t = 0$, assuming that up to that date, all agents had not anticipated the change in future price of investment. Then, at $t = 1$, all agents receive the information of the sequence of current and future tbtc changes and act accordingly, i.e., workers choose occupations according to the new equilibrium sequence $\{v_t^j\}_{t=1}^{\infty}$. Appendix B describes in detail our formulation of the equilibrium conditions of the economy in relative differences from the initial BGP.

Figure 8 shows the evolution of the price of equipment, relative to the price of consumption.

Throughout all the sample, the price of equipment presented a downward trend. Before 1980, the relative price of equipment was falling at a rate of roughly 1.5% per year, while after 1980, it fell at a rate of 3% per year. Since our economy is in a BGP before 1980 and the price of equipment in the model is constant, the specific shock we feed to our model is a decrease in ξ_t^M of 1.5% per year, for 40 years, after which it levels-out. This matches the differential in the evolution of the price of equipment before and after 1980.

Figure 8: Evolution of the price of equipment in the United States (1980-2010)



As we mention in the previous section, parameters ν_j govern the degree of dispersion in the productivity of labor and machines in the production of tasks in the different occupations and are directly linked to the elasticity of substitution of labor and machines in production. Given that we feed a single shock that is common across all occupations, ξ_t^M, the response of occupational employment to this shock in the different occupations will depend on (1) how costly is for workers to switch to other occupations, i.e. the elasticity of labor supply across occupations, which is goverened by different parameters of the worker's problem, and (2) how easy or difficult it is for firms to substitute across different factors of production, which is governed by different parameters of the firm's problem, and in particular, ν_j. Given our shock, we calibrate ν_j to approximate, with the model, the observed changes in the shares of occupational employment between 1980 and 2016.[35]

Table 7 shows the observed changes in employment across the occupations we study, both in the data (column 2) and in the model (column 3). The last column in the table shows our calibrated value for parameters ν_j. A few important patterns emerge. First, the values of ν_j tend to be very different across occupations, ranging between a high elasticity of substitution between labor and machines in production occupations, to much lower values in occupations with a non-routine component. While the aggregate production function in Proposition 3 admits values for ν greater

[35] To estimate the elasticity of substitution between different factors of production, Krusell et al. (2000) also use information on the evolution over time of the share of earnings of different factors. Albeit less formal, our approach to obtain values for ν_j is similar.

than −1, our microfoundation restricts the range of ν_j to be positive. However, the increase in the share of employment of professional and technical occupations over the last few decades was so large that our model, given the only shock we analyze, can only reconcile such an increase with a substantial complementarity between labor and capital in that occupation.[36]

Table 7: Employment changes and elasticity of substitution

Occupations	data	Model	ν_j
Managers	3.1	3.9	0.08
Professionals	8.3	5.8	-0.90
Services	3.7	4.2	0.05
Sales	1.8	1.1	0.10
Office & Admin	-4.9	-3.6	2.00
Construction	-0.6	-0.8	1.00
Repair & Maint	-0.9	-1.1	1.50
Production	-9.1	-8.0	2.50
Transportation	-1.4	-2.3	1.50

Note: changes in the data are between 1980 and 2006. Computed using CPS monthly data for all employed workers between ages 22 and 65 years old.

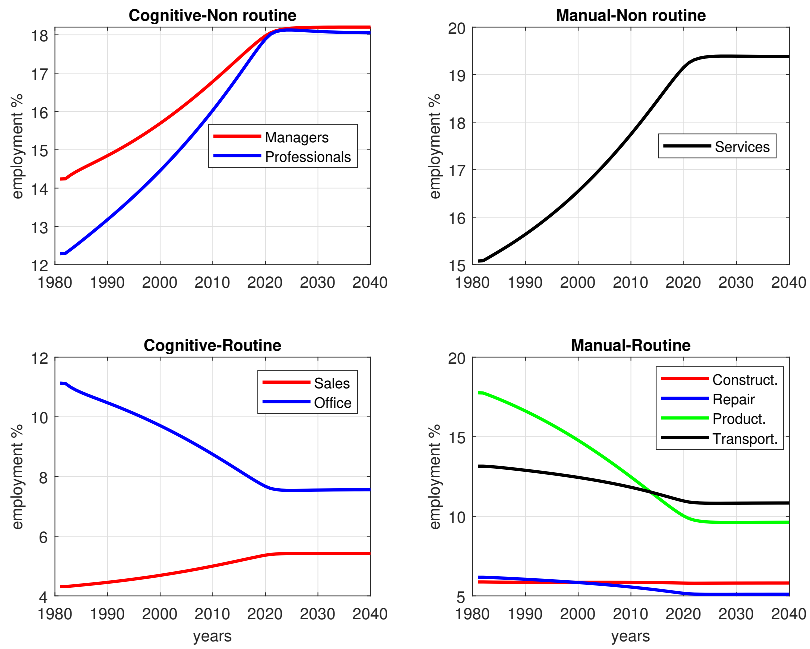
5.3 Aggregate effects of task biased technological change

Through the lens of our model, the data clearly shows that the tbtc shocks had asymmetric effects across the different occupations. Figure 9 shows the dynamic evolution of the employment shares of the four broad occupation groups induced by the calibrated tbtc shock, which complements the information of Table 7. The largest impact is in routine-manual occupations, composed of construction, installation and repair, transportation and production, with a sharp reduction in the employment share of over 12 percentage points, which, as expected, is consistent with the changes observed in the data since 1980 since these are targets of our calibration. Similarly, the combined employment share of routine-cognitive occupations, such as sales and administrative support also falls, but the decrease in this case is more moderate, around 2.5 percentage points.

On the other hand, non-routine occupations, which in the data are typically the polar opposites in terms of wage levels, see their employment shares increase. The increase is more pronounced for the cognitive non-routine occupations such as management, business and professional, with a combined share almost 10 percentage points higher due to the task-biased shock. Non-routine manual occupations such as personal services see their share go up by 4 percentage points due to the same shock.

[36]Alternatively, we could introduce an increase in the productivity of labor in those occupations, however we decided to keep the analysis simple and have only one exogenous shock that is common across all occupations. A different avenue would be to introduce different education groups of workers with an exogenous increase in the supply of college educated cohorts. In the present work we also abstract from this alternative as well, but the model can be extended as we show in Appendix C.

Figure 9: Evolution of employment shares by broad occupation groups



These changes in employment across occupations in the model are the result of two different forces: *(i)* the responses of new cohorts that enter every period, and change their choice of initial occupations; and *(ii)* the response of older workers, as they change their behavior in terms of occupation mobility. In Table 8 we separate the contribution of the two forces, i.e. how much can be accounted by the change in the behavior of new cohorts if older workers do not modify their occupational choices, and how much can be accounted by the change in behavior of the older workers if the new workers do not adapt.

Table 8 shows the initial employment distribution (column 2) and the distribution in the model at period 40, i.e. the end of the tbtc shock (column 3.) In column 4, we account for the accumulated contributions of the responses of the entering cohorts, keeping the behavior of older cohorts as before the shocks. In column 5, we report the contribution of the changes in the mobility patterns of workers once they have entered, i.e. changes in matrix μ. We can see that the bulk of the changes in employment share is the result of the changes in mobility of workers once they have entered labor markets. It is worth noting that those changes do not require a drastically different matrix μ. Small changes in the flows workers between occupations, due to small changes in μ, accumulate over time and lead to substantial changes in the stocks.

The recent literature (Acemoglu & Restrepo, 2018; Karabarbounis & Neiman, 2013) argues that the fall in the labor share in the United States (and around the world) over the last few decades may be driven by technological changes that are biased towards a more intensive use of machines in production. According to the Bureau of Labor Statistics, the labor share of the U.S. nonfarm business sector fluctuated around 0.66 in the 1960s and early 1970s, but since then it

Table 8: Contribution of Responses of Old and New Cohorts to Employment Changes

Occupations	Initial share	Final share	Contribution of Changed Behavior New Cohorts Only	Old Cohorts Only
Managers	14.2	18.1	0.3	3.6
Professionals	12.3	18.0	0.7	5.0
Services	15.1	19.3	0.5	3.7
Sales	4.3	5.4	0.1	1.0
Office & Admin	11.1	7.6	-0.4	-3.2
Construction	5.9	5.8	0.0	-0.1
Repair & Maint	6.2	5.1	0.0	-1.0
Production	17.8	9.9	-1.0	-7.1
Transportation	13.2	10.9	-0.1	-2.1

has trended down towards 0.57 by 2016. This is a 9 percentage point difference.[37] In our model, the drop in the labor share is even more pronounced, around 14 percentage points. As technology reduces the rental rate of machines, a larger share of tasks is performed by machines. This increase is larger in occupations with a higher value of ν_j, as labor and capital are more sustitutable.

In our economy, 40 periods after the shock total output is 75% higher than in the economy with no shock. An important question arises: If machines are receiving a larger share of output, are workers worse-off? Figure 10 provides a partial answer. It shows the evolution of aggregate earnings (left panel) and average earnings (right panel) by broad occupation groups. Total earnings increase across all occupations, and more so in the non-routine groups. However, this is largely the result of higher number of workers in those occupations, as can be inferred by the evolution of average earnings. Importantly, the dispersion in average earnings is more compressed across occupations because workers self-select according to comparative advantage, and tend to stay in their current occupations even if the demand does not increase at the same pace as in other jobs. In general, the expansion of non-routine occupations is driven by workers with weaker comparative advantage there, pushing up the average earnings in the occupations from which they leave and bringing it down in the occupations to which they arrive. This effect will be reflected in earnings inequality as we discuss next.

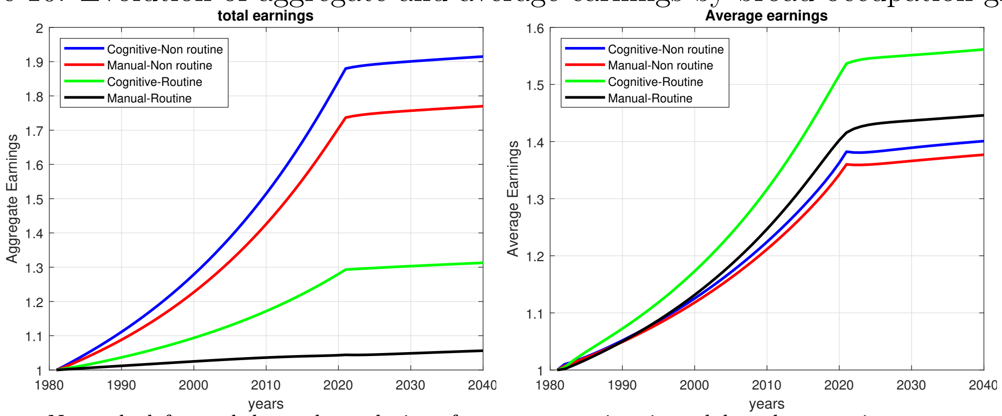
5.4 Effects of Task-biased Innovations on Inequality

We now use the model to assess the impact of task-biased innovations on income distribution. To this end, we simulate histories of individual earnings for a large panel of workers. As technological innovations change unit wages, the ensuing changes in occupational decisions of both new and active workers shape the response of the aggregate and individual evolution of human capital.

As we discussed in Section 3, individual earnings depend on all the history of ϵ shocks and the occupation decision of workers. To compute the history of earnings for individuals, we need

[37] We obtain the data for the labor share from FRED (https://fred.stlouisfed.org/series/PRS85006173). Accessed on October 2019.

Figure 10: Evolution of aggregate and average earnings by broad occupation groups



Note: the left panel shows the evolution of aggregate earnings in each broad occupation group after the tbtc, relative to the initial BGP. Similarly, the right panel shows the evolution of the average earnings.

to recover the parameters λ^j and the matrices τ and χ, and to do so, we use the equilibrium conditions in the initial BGP.[38] Figure 11 shows the recovered level for those parameters, where we assume that the diagonal of the matrix τ, shown in panel (a), is equal to one.[39]

From Figure 11, it is clear that the model requires substantial amount of heterogeneity in order to reconcile the earnings changes and occupational mobility patterns of workers in the initial BGP. First, notice that some occupation switches can be very costly in terms of human capital (e.g.: switching from management to sales) while others are less costly (e.g.: from sales to production or management.) In terms of non-pecuniary terms, recall that with $\gamma = 2$ the values v^j are negative, so a low value is of χ is a positive for utility. Hence, the utility costs of moving from professional occupations to anything else, except perhaps administrative jobs, seem substantial.

Similarly, Table 9 contains information on the scale parameters λ^j, which affect the scale of the distribution of ϵ shocks, with a permanent effect on income and human capital dynamics, and parameters τ^0 and χ^0 affecting the distribution of human capital and employment decision over occupations for the entering cohort.[40]

With this information, we simulate employment histories for workers and compute measures of income inequality.[41] Figure 12 shows the distribution of earnings in the model in the initial at the beginning of the period (1980). The variance of log-earnings in that period is 0.49, consistent with the data. Moreover, the distribution of earnings has a "fat" right tail, which is inherited from the Frechet distribution and the persistent nature of the ϵ shocks and the selection over those and over occupations.

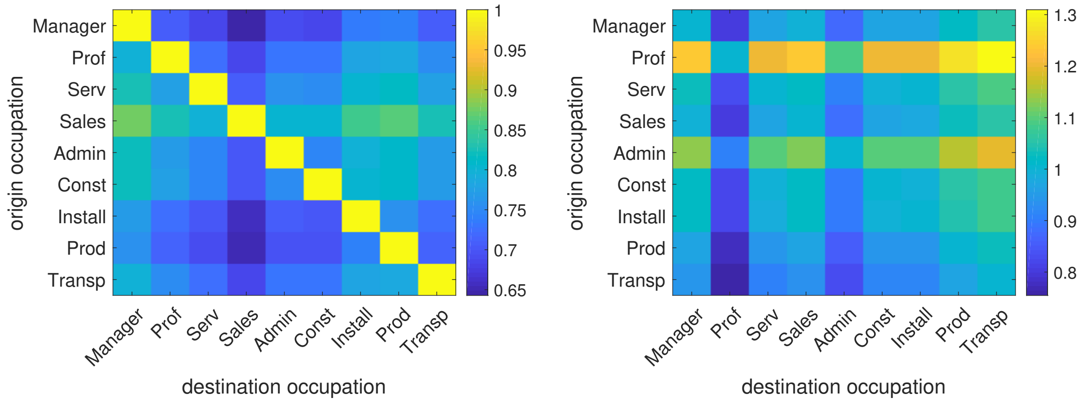
[38] Proposition 4 indicates that it is not necessary to recover the levels of some parameters, particularly λ^j and the matrices τ and χ, to solve for the equilibrium dynamics of aggregate variables neither for the baseline nor counterfactual economies.

[39] We make this normalization since the diagonal of matrix τ cannot be identified separately from the value of the vector λ.

[40] We normalize the first element of χ^0.

[41] We simulate a panel of 40,000 individuals over 200 periods which we use in all our computations in this section.

Figure 11: Human capital transferability and Non-pecuniary mobility costs



(a) human capital transferability matrix (τ) (b) non-pecuniary cost/benefit matrix (χ)

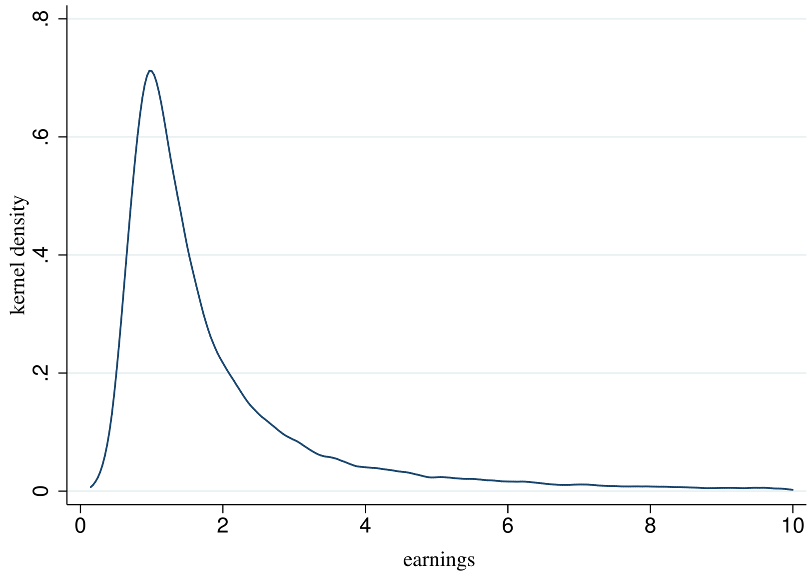
Table 9: Parameters affecting earning dynamics

Occupations	λ	τ^0	χ^0
Managers	0.953	1.043	1.000
Professionals	0.975	0.923	0.914
Services	0.963	0.579	1.675
Sales	0.952	0.806	1.145
Office & Admin	0.967	0.750	1.224
Construction	0.952	0.895	1.104
Repair & Maint	0.948	0.964	1.055
Production	0.952	0.933	1.241
Transportation	0.952	0.824	1.349

Task-biased technical changes affect the evolution of earnings due to changes in wages in different occupations, occupational choices and the dynamics of human capital. Figure 13 shows that during the 40-years episode of tbtc, inequality in the economy almost doubles relative to the economy without the shock. The change in relative terms subsides over time and completely halts some years after the shock stops. In absolute terms, tbtc increase both average earnings and absolute dispersion. Yet, the impact of tbtc goes beyond the absolute levels of incomes, since, as we have emphasized in this paper, the rate at which workers accumulate earnings depends on the endogenous allocation of workers and human capital across occupations. In our calibrated economy, the long-run growth rate in only driven by the exogenous change in the productivity of labor, G_A, which we set at 1%. Yet, the growth rate of human capital over a worker's life cycle does change and after the tbtc shock the rate of human capital accumulation increases. Thus aggregate GDP not only grows due to the aggregate technological change, but also due to an increase in the total accumulation of human capital.

As we discussed in the previous section, the selection of workers over occupations has an effect on inequality. In particular, as the workers with the least strong comparative advantage across all

Figure 12: Initial distribution of earnings in 1980



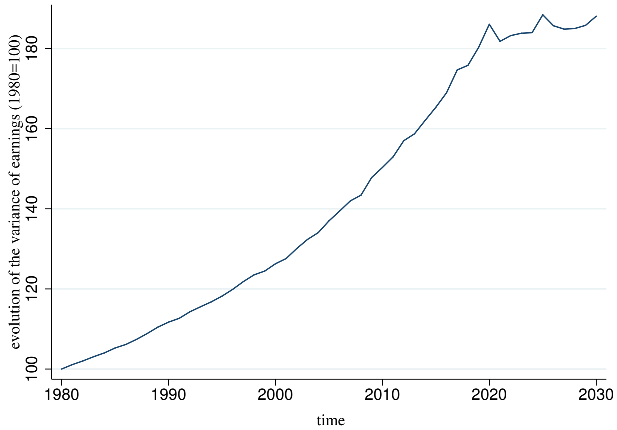
Note: computed using model simulations for the initial period.

occupations are the ones that will respond more to the changes in economic conditions. That is, workers relatively less skilled in routine occupations will tend to switch to non-routine occupations, thus lowering the amount of inequality in routine occupations and increasing it in non-routine ones (in relative terms). Figure 14 shows exactly this pattern. The different panels show the evolution of inequality within each occupation (note that the wage per unit of human capital is the same for workers in each occupation), where panels group occupations in the four broad categories. The top two panels show the evolution of the variance of earnings for cognitive non-routine occupations (top-left) and manual non-routine occupations (top-right). The two panels on the bottom show the same information but for cognitive routine occupations (bottom-left) and manual routine occupations (bottom right). While the effect of the these “more mobile” workers on inequality is not overly large, the figure does show that inequality towards the end of the studied period is higher in non-routine occupations than in routine.

5.5 The role of the elasticity of labor supply and labor demand

In this section we study the importance of labor supply factors and labor demand factors in accounting for job polarization and inequality. Clearly, the effects job polarization and inequality will depend on how easy or difficult is for workers to switch occupations. In an extreme case, if workers cannot switch occupations, we would not see a change in the share of employment across occupations, but most likely inequality would be higher. These are the labor supply factors. On the other hand, if all the different factors of production can easily be substituted in the production process, then all occupations would be affected by tbtc, and in this case we would expect a much

Figure 13: Evolution of earnings inequality



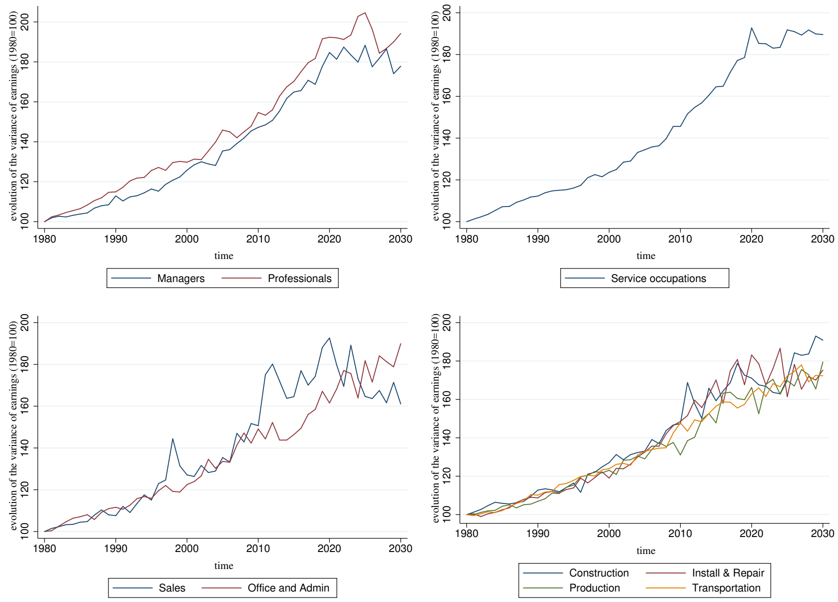
Note: computed using model simulations. The graph shows the difference in the evolution of the variance of earnings between the economy with TBTC and the economy with no change in technology.

lower change in the share of employment across occupations and a lower increase in inequality. These are labor demand forces.

Figure 15 shows the evolution of inequality for our baseline economy and for two counterfactual economies. The first is an economy where all the elasticity of substitution between machines and labor in each occupation j, which is governed by parameter ν^j, is the same. We pick the value of $\nu^j = 0.5$. Relative to the baseline, earnings inequality in this economy is lower for two reasons, on the one hand, the effect of tbtc is symmetric across occupations, and wages in all occupations change by the same magnitude. Because of this, the dynamics of human capital are different, with a lower depreciation of human capital due to the lower amount of reallocation, and also, to a slower pace of accumulation of human capital, which tends to occur in routine occupations. Note that, in this economy, there is no job polarization as the effects of technology is symmetric across different occupations leading to no net changes in employment for the different occupations.

The second case we explore is an economy where we restrict workers from moving. In our exercise, we keep the evolution of wages identical as in the baseline economy and focus only on the effects of human capital accumulation. The first effect, which cannot be seen in the figure given that all economies are normalized by an economy with no tbtc shock, is that if we restrict mobility, the overall accumulation of human capital is much lower. In this case, we prevent workers from selecting their most favourable comparative advantage shocks ϵ, which on average translate into a lower human capital accumulation. The second pattern is that inequality in this economy, is larger than in the baseline case. In this economy, workers cannot "escape" the losses from automation nor move into jobs for which they have a better productivity. The compression of earnings due to self-selection that is typically present in Roy models is not operative in this case, which leads to a higher level of inequality. Note that, similarly to the previous case, as workers cannot move, the

Figure 14: Inequality across occupation groups



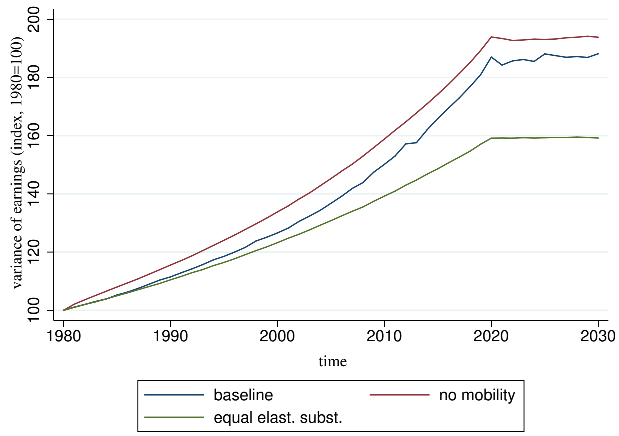
Note: computed using model simulations. The graph shows the difference in the evolution of the variance of earnings between the economy with TBTC and the economy with no change in technology.

economy does not exhibit job polarization either.

Clearly, in an economy where workers can reallocate across occupations with no frictions, inequality will be very compressed.

Overall, these exercises suggest that supply factors, which in our case mean mobility frictions, tend to exacerbate earnings inequality, but reduce the amount of reallocation across occupations and polarization. On the other hands, demand factors, like a more similar way to substitute factors of production in the different occupations, tend to reduce the amount of inequality, but also lead to a lower labor reallocation across occupations, and thus a lower job polarization.

Figure 15: Evolution of earnings inequality



Note: computed using model simulations. The graph shows the difference in the evolution of the variance of earnings between the economy with TBTC and the economy with no change in technology. Baseline case is our baseline economy. No mobility where individuals cannot switch occupations. Equal elasticity of substitution is an economy where parameter ν^j affecting the substitution possibilities across machines and labor in different economies is the same and equal to 0.5.

6 Conclusion

We develop a dynamic Roy model of occupational choice with human capital accumulation and use it to explore the general equilibrium effects of new technologies on the labor market. In our model, infinitely-lived workers can switch occupations in any period to maximize their lifetime utility. In our setting, a worker's human capital is driven by his labor market choices, given idiosyncratic occupation-specific productivity shocks and the costs of switching occupations. We first characterize the equilibrium assignment of workers to jobs. A key result is that the resulting evolution of aggregate human capital across occupations ultimately determines the long-run rate of growth of the economy. We then use the model to quantitatively study how worker's individual occupation choices change with the introduction of new technologies, and in turn how this choices shape the equilibrium allocation of workers to different jobs, the dynamics of aggregate human capital, the behavior of earnings inequality, the evolution of the labor share, and the welfare of the different workers in the economy.

The paper has a number of methodological contributions. First, we fully characterize the solution of the recursive problem of a worker under standard CRRA preferences when the worker is subject to a large number of labor market opportunities shocks in every period affecting her comparative advantage in different occupations. Thus, we bridge recent quantitative work that uses static assignment Roy models with extreme-value shocks with the standard recursive models for households in macroeconomics. In this way, our model generates transition probabilities across occupations over time. Second, we fully characterize the asymptotic behavior of aggregate economies implied by the individual dynamic occupation choices of workers. For any given vector of skill prices, we show that the economy converges to a unique invariant distribution of workers. Although the Roy model has been studied and used in great length, we uncover important new features which are present only in a dynamic context. We show that, generically, the reallocation of workers to occupations combined with the accumulation of occupational human capital leads to sustained growth over time for the economy. The growth rate in our model is endogenously determined by the equilibrium occupational choices, and thus, changes in economic conditions that alter worker's choices affect the long-run growth rate of the economy. Third, we embedded the workers' problem in a fairly rich general equilibrium environment where different types of workers are allocated to different tasks in production. We derive a very transparent and tractable aggregation that arises from the assignment of workers to tasks. Then, we show the existence of a competitive-equilibrium balanced-growth path, and for a simple version of our model we can also characterize uniqueness. Fourth, by incorporating two forms of physical capital, we provide a quantitative framework to study the impact of automation and other task biased technological improvements on the earnings of different occupations. Our model of production and tasks generates an intuitive expression that directly links the labor share of the economy with wages, rental rates and the productivity of different types of labor and capital, allowing us to study the effects of technology on the labor share of the economy. Fifth, we extend recent dynamic-hat-algebra methods and show they can be used with more general preferences (CRRA) and with human capital accumulation. As with other hat-algebra methods, the advantage is a substantially reduced set of calibrated parameters needed for the quantitative application of the model. Sixth, we discuss a variety of relevant extensions of our baseline model, ranging from workers' age and ex-ante heterogeneity, endogenous on-the-job training and occupation-specific automation.

Using our model we make a number of substantial contributions. Mapping our model to the moments observed in the 1970s for the U.S. economy, we account for the changes in employment across occupations and the increase in earnings inequality that arise from task biased technological advances. An important change observed in U.S. labor markets in the past few decades is the polarization of skills in the labor market. That is, the decline of employment in middle-skill occupations, like manufacturing and production occupations, and the growth of employment in both high and low-skill occupations, like managers and professional occupations on one end, and assisting or caring for others on the other. Using our model we show how some task biased technical improvements can jointly explain the increase in polarization, earnings inequality and occupational mobility in U.S. labor markets.

In addition, our dynamic model highlights the long-lasting impact of permanent, but once-and-for-all technological changes. Indeed, in our dynamic setting, once-and-for-all changes in automation or other technological changes can lead to sustained growth effects. Our quantitative exercise highlight how this growth effect changes the conclusion on earnings inequality and welfare. We emphasize that the welfare and inequality implications for technological changes can be vastly richer than those obtained in other settings as they originate not only from changes in skills prices in each period but also on changes in the equilibrium growth rate of earnings. Thus, on the one hand, the positive impact on some workers is not only due to higher level of earnings but also from a faster growth. On the other hand, some workers can be worse-off due to lower levels of earnings and a higher rate at which they change occupations. These aspects are fully incorporated in our exercises.

Our theory opens a number of exciting avenues for future research. International trade, offshoring, and migration policies, can alter the demand for workers in different occupations in an asymmetric fashion. Our model highlights how these type of changes affect not only the distribution of workers and human capital in different occupations, but also affect the rate at which workers accumulate human capital over their life cycle, and depending on parameters, may affect the long-run growth of the economy.

References

Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics* (Vol. 4, pp. 1043–1171). Elsevier.

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

Acemoglu, D., & Restrepo, P. (2019). Robots and jobs: Evidence from U.S. labor markets. *Journal of Political Economy, forthcoming*.

Adao, R., Beraja, M., & Pandalai-Nayar, N. (2018). Skill-biased technological transitions. *Working Paper*.

Alvarez, F., & Stokey, N. L. (1998). Dynamic programming with homogeneous functions. *Journal of economic theory*, *82*(1), 167–189.

Autor, D., & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the U.S. labor market. *American Economic Review*, *103*(5), 1553–97.

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

Autor, D., Levy, F., & Murnane, R. J. (2003, 11). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, *118*(4), 1279-1333.

Burstein, A., Morales, E., & Vogel, J. (2018). Changes in between-group inequality: computers, occupations, and international trade. *American Economic Journal: Macroeconomics*.

Caliendo, L., Dvorkin, M., & Parro, F. (2019). Trade and labor market dynamics: General equilibrium analysis of the China trade shock. *Econometrica*, *87*(3), 741–835.

Cortes, M., Nekarda, C., Jaimovich, N., & Siu, H. (2016). The micro and macro of disappearing routine jobs: A flows approach. *Working Paper*.

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

Doms, M., & Lewis, E. (2006). Labor supply and personal computer adoption. *Federal Reserve Bank of Philadelphia, Working Paper*.

Eaton, J., & Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, *70*(5), 1741–1779.

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

Galle, S., Rodríguez-Clare, A., & Yi, M. (2017). Slicing the pie: Quantifying the aggregate and distributional effects of trade. *NBER Working Paper*.

Gantmacher, F. R. (1959). *The theory of matrices.* NY: Clelsea.

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

Goos, M., Manning, A., & Salomons, A. (2014, August). Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review*, *104*(8), 2509-26.

Greenwood, J., Hercowitz, Z., & Krusell, P. (1997). Long-run implications of investment-specific technological change. *American Economic Review*, *87*(3), 342-62.

Guvenen, F., Kuruscu, B., Tanaka, S., & Wiczer, D. (2019). The micro and macro of disappearing routine jobs: A flows approach. *Working Paper*.

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

Hsieh, C.-T., Hurst, E., Jones, C. I., & Klenow, P. J. (2019). The allocation of talent and us economic growth. *Econometrica*, *87*(5), 1439–1474.

Kambourov, G., & Manovskii, I. (2008). Rising occupational and industry mobility in the United States: 1968–97. *International Economic Review*, *49*(1), 41–79.

Kambourov, G., & Manovskii, I. (2009). Occupational mobility and wage inequality. *The Review of Economic Studies*, *76*(2), 731–759.

Kambourov, G., & Manovskii, I. (2013). A cautionary note on using (March) Current Population Survey and Panel Study of Income Dynamics data to study worker mobility. *Macroeconomic Dynamics*, *17*(1), 172–194.

Karabarbounis, L., & Neiman, B. (2013). The global decline of the labor share. *The Quarterly Journal of Economics*, *129*(1), 61–103.

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

Lagakos, D., Moll, B., Porzio, T., Qian, N., & Schoellman, T. (2018). Life cycle wage growth across countries. *Journal of Political Economy*, *126*(2), 797-849.

Lagakos, D., & Waugh, M. E. (2013). Selection, agriculture, and cross-country productivity differences. *American Economic Review*, *103*(2), 948–80.

Lee, D., & Wolpin, K. I. (2006). Intersectoral labor mobility and the growth of the service sector. *Econometrica*, *74*(1), 1-46.

Lillard, L. A., & Willis, R. J. (1978). Dynamic aspects of earning mobility. *Econometrica*, 985–1012.

Lucas, R. E., & Moll, B. (2014). Knowledge growth and the allocation of time. *Journal of Political Economy*, *122*(1), 1–51.

Luttmer, E. G. (2007). Selection, growth, and the size distribution of firms. *The Quarterly Journal of Economics*, *122*(3), 1103–1144.

MaCurdy, T. E. (1982). The use of time series processes to model the error structure of earnings in a longitudinal data analysis. *Journal of Econometrics*, *18*(1), 83–114.

Martinez, J. (2019). Automation, growth and factor shares. *Working Paper*.

Silva, J. S., & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, *88*(4), 641–658.

Stokey, N., Lucas Jr, R., & Prescott, E. (1989). *Recursive methods in economic dynamics*. Harvard University Press.

A Proofs

We assume that $0 < \underline{w} \equiv \min_j \{w^j\} < \bar{w} \equiv \max_j \{w^j\} < +\infty$. As in the text, denote the flow utilities as $u^j \equiv \Upsilon^j \frac{(w^j)^{1-\gamma}}{1-\gamma}$ for $\gamma \neq 1$ and $u^j \equiv \Upsilon^j + \log(w^j)$ for $\gamma = 1$. Also, denote $\underline{u} = \min_j u^j$ and $\bar{u} = \max_j u^j$. For all values of $\gamma \geq 0$, the verify conditions on $\tau_{j,\ell}$ and ϵ_ℓ such that we ensure the existence of upper and lower bounds $-\infty < \underline{v} < \bar{v} < +\infty$, such that the relevant range of normalized conditional expected values remain within those bounds, i.e. $v \in [\underline{v}, \bar{v}]^J$. The bounds $\underline{v}$, $\bar{v}$ depend on the utility curvature γ as we explain in detail below.

Proof. Lemma 1: Existence and Uniqueness BE. For any vector $v \in R^J$, define the operator $T : R^J \to R^J$ so that, for each entry j:

$$(Tv)_j = u^j + \beta\, E_\epsilon \max_\ell \left\{ v^\ell \ (\tau_{j\ell}\epsilon_\ell)^{1-\gamma} \right\}.$$

Consider first the case $0 \leq \gamma < 1$. Here, $u^j > 0$, and the relevant values of v are all positive. Denote also $\bar{\Phi} = \max_j \Phi_j$. In this case, we can define $\underline{v} = 0_J$, the $J \times 1$ vector of zeros, as a lower-bound. As an upper-bound, let $\bar{v} = 1_J \times \frac{\bar{u}}{1-\beta\bar{\Phi}}$, i.e. the vector with its J entries equal to $\frac{\bar{u}}{1-\beta\bar{\Phi}}$. We first confirm these bounds. Start by noticing that T is monotone, i.e., if $v_1 \geq v_2$ (pairwise,) then $Tv_1 \geq Tv_2$. Next, notice that $T\bar{v} \leq \bar{v}$, since, for all j:

$$\begin{aligned}(T\bar{v})_j &= u^j + \beta\, E_\epsilon \max_\ell \left\{ \bar{v}\, (\tau_{j\ell}\epsilon_\ell)^{1-\gamma} \right\} \\ &\leq \bar{u} + \beta\Phi_j \bar{v} = \frac{\bar{u}}{1-\beta\bar{\Phi}}.\end{aligned}$$

Finally, notice that, $(T0_J) > 0_J$, since for any j:

$$(T0_J)_j = u^j + \beta\, E_\epsilon \max_\ell \left\{ 0 \times (\tau_{j\ell}\epsilon_\ell)^{1-\gamma} \right\} = u^j > 0.$$

This proves that for any $v \in [0_J, \bar{v}]$ then $Tv \in [0_J, \bar{v}]$. Let any v_1,$v_2 \in [0_J, \bar{v}]$. Then

$$\begin{aligned}\|Tv_1 - Tv_2\| &= \max_j \left| u^j + \beta\, E_\epsilon \max_\ell \left\{ v_{1,\ell} \ (\tau_{j\ell}\epsilon_\ell)^{1-\gamma} \right\} - u^j - \beta\, E_\epsilon \max_\ell \left\{ v_{2,\ell} \ (\tau_{j\ell}\epsilon_\ell)^{1-\gamma} \right\} \right| \\ &= \beta \max_j \left| \ E_\epsilon \max_\ell \left\{ v_{1,\ell} \ (\tau_{j\ell}\epsilon_\ell)^{1-\gamma} \right\} - \ E_\epsilon \max_\ell \left\{ v_{2,\ell} \ (\tau_{j\ell}\epsilon_\ell)^{1-\gamma} \right\} \right| \\ &\leq \beta \max_j \left| \ E_\epsilon \max_\ell \left\{ \ (\tau_{j\ell}\epsilon_\ell)^{1-\gamma} \right\} \right| \max_\ell |\, v_{1,\ell} - v_{2,\ell}| \\ &= \beta\bar{\Phi}\, \|v_1 - v_2\|\,.\end{aligned}$$

This verifies that T is a contraction when $\beta\bar{\Phi} < 1$. Using the distance function induced by the uniform norm, $\|v\| = \max_j |v^j|$, the space $([0_J, \bar{v}]\,, \|\cdot\|)$ is a complete metric space, and hence a unique fixed point of T exists and is unique.

We next consider the special, case of $\gamma = 1$. In this case, $u^j \equiv \ln w^j$, and the Bellman equation (4) becomes

$$V\,(j, h, \epsilon) = u^j + \beta \max_\ell \left\{ E_{\epsilon'} V\,[\ell,\, h',\, \epsilon'] \right\}. \tag{43}$$

Following the same steps as when $\gamma < 1$, we now verify the hypothesis that $E_\epsilon V\,[j,\, h,\, \epsilon] = v^j + B \ln h$, for some B. Taking expectations in both sides of (43), and using the law of motion for the accumulation of human capital (3), equation (43) becomes:

$$v^j + B \ln h = u^j + \beta E_\epsilon \left[\max_\ell \left\{ v^\ell + B\,[\ln(h) + \ln(\tau_{j,\ell}\epsilon_\ell)] \right\} \right].$$

Then, the hypothesis is confirmed with $B = 1/\,(1-\beta)$, and v^j defined as a fixed point of the operator $T : R^J \to R^J$

so that, for each entry j:

$$(Tv)_j = u^j + \beta E_\epsilon \left[\max_\ell \left\{ v^\ell + \frac{\ln(\tau_{j,\ell}\epsilon_\ell)}{1-\beta} \right\} \right]. \tag{44}$$

Conditions for existence and uniqueness of a fixed point $v \in \mathbb{R}^J$ are as follows. We first ensure that there are finite lower and upper bounds. To this end, notice that the distribution of ϵ_ℓ is such that $-\infty < E_\epsilon [\max_\ell \{\ln(\tau_{j,\ell}\epsilon_\ell)\}] < \infty$. Then, the overly slack bounds can be set to:

$$\begin{aligned} \underline{v} &= \frac{\ln \underline{w} + \beta \left(\frac{\ln(\tau_{\min})}{1-\beta} + \frac{E_\epsilon[\max_\ell\{\ln(\epsilon_\ell)\}]}{1-\beta} \right)}{1-\beta}, \\ \bar{v} &= \frac{\ln \bar{w} + \beta \left(\frac{\ln(\tau_{\max})}{1-\beta} + \frac{E_\epsilon[\max_\ell\{\ln(\epsilon_\ell)\}]}{1-\beta} \right)}{1-\beta}. \end{aligned}$$

Clearly, these bounds are overly slack and both are finite as long as $\tau_{\min} \equiv \min_{j,\ell} \{\tau_{j,\ell}\} > 0$ and $\tau_{\max} \equiv \max \{\tau_{j,\ell}\} < \infty$, which we have assumed all along. It is straightforward to verify that the operator T defined by (44) maps the rectangle $[\underline{v}, \bar{v}]$ into itself and it is monotone. It is also straightforward to verify that as long as $0 < \beta < 1$, then T is a contraction on the complete metric space $([\underline{v}, \bar{v}], \|\cdot\|)$.

Finally, consider the case of $\gamma > 1$. Here, all $u^j < 0$, and the relevant values of v are all negative. In this case, we can define the upper-bound to be $\bar{v} = 0_J$, i.e. the $J \times 1$ vector of zeroes; the lower-bound can be taken to be $\underline{v} = 1_J \times \frac{u}{1-\beta\bar{\Phi}}$, i.e. the vector with its J entries equal to $\frac{u}{1-\beta\bar{\Phi}}$. To confirm these bounds, first notice that for any $v \le 0$, $Tv < 0_J$. In particular, $(T0_J) < 0_J$, since for any j:

$$(T0_J)_j = u^j + \beta\, E_\epsilon \max_\ell \left\{ 0 \times (\tau_{j\ell}\epsilon_\ell)^{1-\gamma} \right\} = u^j < 0.$$

Next, notice that $T\underline{v} \ge \underline{v}$, since, under the definition of $\Phi_j = E_\epsilon \min_\ell \left\{ (\tau_{j\ell}\epsilon_\ell)^{1-\gamma} \right\}$ and the assumption that $\beta\bar{\Phi} < 1$, for all j:

$$\begin{aligned} (T\underline{v})_j &= u^j + \beta\, E_\epsilon \max_\ell \left\{ \underline{v}\, (\tau_{j\ell}\epsilon_\ell)^{1-\gamma} \right\} \\ &= u^j + \beta\, E_\epsilon \max_\ell \left\{ (-\underline{v})\, (-\tau_{j\ell}\epsilon_\ell)^{1-\gamma} \right\} \\ &= u^j - \beta\, E_\epsilon \min_\ell \left\{ (\tau_{j\ell}\epsilon_\ell)^{1-\gamma} \right\} (-\underline{v}) \\ &= \bar{u} + \beta \Phi_j \underline{v} \\ &\ge \bar{u} + \beta\bar{\Phi}\underline{v} = \frac{\bar{u}}{1-\beta\bar{\Phi}}. \end{aligned}$$

The operator T is obviously monotone and the condition $\beta\bar{\Phi} < 1$ ensures that it satisfies discounting. ∎

Proof. Lemma 2: Derived Distributions: Write $F_X(x) = \Pr[X \le x]$, where $x > 0$. Consider first the case $\gamma < 1$. Then:

$$\begin{aligned} F_X(x) &= \Pr[X \le x] = \Pr\left[\epsilon^{1-\gamma} \le x\right] = \Pr\left[\epsilon \le (x)^{\frac{1}{1-\gamma}}\right] \\ &= F_\epsilon\left((x)^{\frac{1}{1-\gamma}}\right) = e^{-\left(\frac{(x)^{\frac{1}{1-\gamma}}}{\lambda_\ell}\right)^{-\alpha}} = e^{-\left(\frac{x}{(\lambda_\ell)^{1-\gamma}}\right)^{-\left[\frac{\alpha}{1-\gamma}\right]}}, \end{aligned}$$

which is the cdf of a Frechet distribution with scale $(\lambda_\ell)^{1-\gamma} > 0$ and curvature $\frac{\alpha}{1-\gamma} > 1$. Consider now the case where $\gamma > 1$. Then,

$$\begin{aligned} F_X(x) &= \Pr[X \le x] = \Pr\left[\epsilon^{1-\gamma} \le x\right] = \Pr\left[\epsilon \ge (x)^{\frac{1}{1-\gamma}}\right] \\ &= 1 - \Pr\left[\epsilon < (x)^{\frac{1}{1-\gamma}}\right] = 1 - e^{-\left(\left(\frac{x}{\lambda_\ell}\right)^{\frac{1}{1-\gamma}}\right)^{-\alpha}} = 1 - e^{-\left(\frac{x}{\lambda_\ell}\right)^{\frac{\alpha}{\gamma-1}}}, \end{aligned}$$

which is the c.d.f. of a Weibull distribution with scale $(\lambda_\ell)^{\gamma-1} > 0$ and curvature $\frac{\alpha}{\gamma-1} > 0$. Finally, consider the case of $\gamma = 1$. Here x can be positive or negative. Anyway,

$$\begin{aligned} F_X(x) &= \Pr[X \le x] = \Pr[\ln \epsilon \le x] = \Pr[\epsilon \le e^x] \\ &= e^{-\left(\frac{e^x}{\lambda_\ell}\right)^{-\alpha}} = e^{-e^{-\frac{[x-\ln\lambda_\ell]}{1/\alpha}}}, \end{aligned}$$

which is the cdf of a Gumbel with location parameter $\ln(\lambda_\ell)$ and shape parameter $1/\alpha > 0$. ■

Proof. Theorem 1: Worker's Optimization Problems. We first prove the statement about the functional form adopted by the value function that solves the BE. Second, we prove the statement on the functional form for the switching probabilities. We skip the proof for existence and uniqueness since it follows from Lemma 1 as a special case.

I. Functional Forms of the BE. As in the text, for all $0 \le \gamma \ne 1$, define

$$\Omega_{j\ell} = \tau_{j\ell}^{(1-\gamma)} v^\ell.$$

We now solve for the expression $E_\epsilon\left[\max_\ell\left\{\Omega_{j\ell}(\epsilon_\ell)^{1-\gamma}\right\}\right]$ and plug the solution in the formula for the value function v^j:

$$\begin{aligned} & E_\epsilon\left[\max_\ell\left\{\Omega_{j\ell}(\epsilon_\ell)^{1-\gamma}\right\}\right] \\ = & \sum_{\ell=1}^{J}\int_0^\infty \Omega_{j\ell}(\epsilon_\ell)^{1-\gamma}\prod_{k\ne\ell}\Pr\left[\epsilon_k < \left(\frac{\Omega_{j\ell}}{\Omega_{jk}}\right)^{\frac{1}{1-\gamma}}\epsilon_\ell\right] f(\epsilon_\ell)d\epsilon_\ell \\ = & \sum_{\ell=1}^{J}\int_0^\infty \Omega_{j\ell}(\epsilon_\ell)^{1-\gamma}\prod_{k\ne\ell} F\left[\left(\frac{\Omega_{j\ell}}{\Omega_{jk}}\right)^{\frac{1}{1-\gamma}}\epsilon_\ell\right] f(\epsilon_\ell)d\epsilon_\ell \\ = & \sum_{\ell=1}^{J}\int_0^\infty \Omega_{j\ell}(\epsilon_\ell)^{1-\gamma}\prod_{k\ne\ell} e^{-\left[\left(\frac{\Omega_{j\ell}}{\Omega_{jk}}\right)^{\frac{1}{1-\gamma}}\frac{\epsilon_\ell}{\lambda_k}\right]^{-\alpha}} e^{-\left(\frac{\epsilon_\ell}{\lambda_\ell}\right)^{-\alpha}}\left(\frac{\alpha}{\lambda_\ell}\right)\left(\frac{\epsilon_\ell}{\lambda_\ell}\right)^{-1-\alpha} d\epsilon_\ell \\ = & \sum_{\ell=1}^{J}\int_0^\infty \Omega_{j\ell}(\epsilon_\ell)^{1-\gamma} e^{-\left[\sum_{k=1}^{J}\left(\left(\frac{\Omega_{j\ell}}{\Omega_{jk}}\right)^{\frac{1}{1-\gamma}}\frac{1}{\lambda_k}\right)^{-\alpha}\right](\epsilon_\ell)^{-\alpha}} \frac{\alpha}{\lambda_\ell}\left(\frac{\epsilon_\ell}{\lambda_\ell}\right)^{-1-\alpha} d\epsilon_\ell \\ = & \sum_{\ell=1}^{J}\int_0^\infty \Omega_{j\ell}(\lambda_\ell)^{\alpha} e^{-\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}(\lambda_k)^\alpha\right](\epsilon_\ell)^{-\alpha}} \alpha(\epsilon_\ell)^{-\gamma-\alpha} d\epsilon_\ell. \end{aligned}$$

The first equality is by definition, the second holds since the ϵ shocks are i.i.d. and therefore the joint c.d.f. can be written as the product of the marginals. The third, fourth and fifth equality hold by using the definition of the c.d.f.s; the sixth groups all terms, and the last two steps result from regrouping and simplifying terms. Now, to solve the resulting integral, change variable and define

$$z \equiv \left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}(\lambda_k)^\alpha\right](\epsilon_\ell)^{-\alpha}.$$

Therefore,

$$\epsilon_\ell = [z]^{\frac{-1}{\alpha}}\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}(\lambda_k)^\alpha\right]^{\frac{1}{\alpha}} \text{ and } d\epsilon_\ell = \frac{-1}{\alpha}[z]^{\frac{-1}{\alpha}-1}\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}(\lambda_k)^\alpha\right]^{\frac{1}{\alpha}} dz.$$

Substituting these three formulas in the last expression, we get:

$$
\begin{aligned}
& E_{\epsilon}\left[\max_{\ell}\left\{\Omega_{j\ell}\left(\epsilon_{\ell}\right)^{1-\gamma}\right\}\right] \\
= & \sum_{\ell=1}^{J}\int_{0}^{\infty}\Omega_{j\ell}\left(\lambda_{\ell}\right)^{\alpha}e^{-\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]\left(\epsilon_{\ell}\right)^{-\alpha}}\alpha\left(\epsilon_{\ell}\right)^{-\gamma-\alpha}\left(d\epsilon_{\ell}\right) \qquad (45)\\
= & \sum_{\ell=1}^{J}\int_{0}^{\infty}\Omega_{j\ell}\left(\lambda_{\ell}\right)^{\alpha}e^{-z}\left(\left[z\right]^{\frac{-1}{\alpha}}\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]^{\frac{1}{\alpha}}\right)^{-\gamma-\alpha}\left[z\right]^{\frac{-1}{\alpha}-1}\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]^{\frac{1}{\alpha}}dz \\
= & \sum_{\ell=1}^{J}\Omega_{j\ell}\left(\lambda_{\ell}\right)^{\alpha}\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]^{\frac{-(\gamma+\alpha)}{\alpha}}\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]^{\frac{1}{\alpha}}\int_{0}^{\infty}e^{-z}\left[z\right]^{-\frac{(1-\gamma)}{\alpha}}dz \\
= & \Gamma\left(1-\frac{1-\gamma}{\alpha}\right)\sum_{\ell=1}^{J}\Omega_{j\ell}\left(\lambda_{\ell}\right)^{\alpha}\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]^{\frac{1-(\gamma+\alpha)}{\alpha}}, \qquad (46)
\end{aligned}
$$

where the last line uses the definition of the gamma function. This expression is valid for all $0 \leq \gamma \neq 1$.

Now, consider the cases when $0 \leq \gamma < 1$, and $\Omega_{jk} > 0$. Since they are positive, we can directly take the terms $\Omega_{j\ell}$ as a common factors from the summation within brackets of (46) we get

$$
\begin{aligned}
E_{\epsilon}\left[\max_{\ell}\left\{\Omega_{j\ell}\left(\epsilon_{\ell}\right)^{1-\gamma}\right\}\right] & = \Gamma\left(1-\frac{1-\gamma}{\alpha}\right)\sum_{\ell=1}^{J}\Omega_{j\ell}\left(\lambda_{\ell}\right)^{\alpha}\left[\left(\Omega_{j\ell}\right)^{\frac{-\alpha}{1-\gamma}}\sum_{k=1}^{J}\left(\Omega_{jk}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]^{\frac{1-(\gamma+\alpha)}{\alpha}} \\
& = \Gamma\left(1-\frac{1-\gamma}{\alpha}\right)\sum_{\ell=1}^{J}\left(\lambda_{\ell}\right)^{\alpha}\left(\Omega_{j\ell}\right)^{\frac{\alpha}{1-\gamma}}\left[\sum_{k=1}^{J}\left(\Omega_{jk}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]^{\frac{1-(\gamma+\alpha)}{\alpha}} \\
& = \Gamma\left(1-\frac{1-\gamma}{\alpha}\right)\left[\sum_{\ell=1}^{J}\left(v^{\ell}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{\ell}\tau_{j\ell}\right)^{\alpha}\right]^{\frac{1-\gamma}{\alpha}},
\end{aligned}
$$

where we have used $\Omega_{j\ell} = \tau_{j\ell}^{(1-\gamma)}v^{\ell}$, and, as claimed in the text

$$
v^{j} = u^{j} + \beta\,\Gamma\left(1-\frac{1-\gamma}{\alpha}\right)\left[\sum_{\ell=1}^{J}\left(v^{\ell}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{\ell}\tau_{j\ell}\right)^{\alpha}\right]^{\frac{1-\gamma}{\alpha}}.
$$

Next, consider $\gamma > 1$. In this case, both v^j and Ω_{jk} are negative. Obviously, we can re-write expression (46) as

$$
E_{\epsilon}\left[\max_{\ell}\left\{\Omega_{j\ell}\left(\epsilon_{\ell}\right)^{1-\gamma}\right\}\right] = -\Gamma\left(1-\frac{1-\gamma}{\alpha}\right)\sum_{\ell=1}^{J}\left|\Omega_{j\ell}\right|\left(\lambda_{\ell}\right)^{\alpha}\left[\sum_{k=1}^{J}\left(\frac{\left|\Omega_{jk}\right|}{\left|\Omega_{j\ell}\right|}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]^{\frac{1-(\gamma+\alpha)}{\alpha}},
$$

where $|\cdot|$ is just the absolute value function. Then, by following exactly the same steps for $\gamma < 1$, we obtain for $\gamma > 1$ the expression

$$
v^{j} = u^{j} - \beta\,\Gamma\left(1-\frac{1-\gamma}{\alpha}\right)\left[\sum_{\ell=1}^{J}\left(-v^{\ell}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{\ell}\tau_{j\ell}\right)^{\alpha}\right]^{\frac{1-\gamma}{\alpha}},
$$

where we have used the fact that $\tau_{j\ell} > 0$ for all j, ℓ.

II. Switching Probabilities. Next, we characterize $\mu_{j\ell}$, the probability that workers attached to occupation j

switch to occupation ℓ. As before, consider first the case of $\gamma < 1$. Then

$$\begin{aligned}
\mu_{j\ell} &= \Pr\left[\Omega_{jk}(\epsilon_k)^{1-\gamma} < \Omega_{j\ell}(\epsilon_\ell)^{1-\gamma}\right] \text{ for all } k \neq \ell \\
&= \Pr\left[(\epsilon_k)^{1-\gamma} < \frac{\Omega_{j\ell}}{\Omega_{jk}}(\epsilon_\ell)^{1-\gamma}\right] \text{ for all } k \neq \ell \\
&= \Pr\left[\epsilon_k < \left(\frac{\Omega_{j\ell}}{\Omega_{jk}}\right)^{\frac{1}{1-\gamma}}\epsilon_\ell\right] \text{ for all } k \neq \ell,
\end{aligned}$$

i.e. the direction of the inequality is preserved both times.The first time because we are dividing by $\Omega_{jk} > 0$ and the second because the transformation $(\cdot)^{\frac{1}{1-\gamma}}$ is monotonically increasing. Consider next the case when $\gamma > 1$. Then

$$\begin{aligned}
\mu_{j\ell} &= \Pr\left[\Omega_{jk}(\epsilon_k)^{1-\gamma} < \Omega_{j\ell}(\epsilon_\ell)^{1-\gamma}\right] \text{ for all } k \neq \ell \\
&= \Pr\left[(\epsilon_k)^{1-\gamma} > \frac{\Omega_{j\ell}}{\Omega_{jk}}(\epsilon_\ell)^{1-\gamma}\right] \text{ for all } k \neq \ell \\
&= \Pr\left[\epsilon_k < \left(\frac{\Omega_{j\ell}}{\Omega_{jk}}\right)^{\frac{1}{1-\gamma}}\epsilon_\ell\right] \text{ for all } k \neq \ell.
\end{aligned}$$

i.e. the direction of the inequalities inside the probabilities is reversed twice: The first time as we are dividing by a negative number, $\Omega_{jk} < 0$, and the second because the transformation $(\cdot)^{\frac{1}{1-\gamma}}$ is monotonically decreasing.

Therefore, for both instances

$$\begin{aligned}
\mu_{j\ell} &= \int_0^\infty \prod_{k\neq\ell} \Pr\left[\epsilon_k < \left(\frac{\Omega_{j\ell}}{\Omega_{jk}}\right)^{\frac{1}{1-\gamma}}\epsilon_\ell\right] f_\ell(\epsilon_\ell)d\epsilon_\ell \\
&= \int_0^\infty \prod_{k\neq\ell} F_k\left[\left(\frac{\Omega_{j\ell}}{\Omega_{jk}}\right)^{\frac{1}{1-\gamma}}\epsilon_\ell\right] f_\ell(\epsilon_\ell)d\epsilon_\ell \\
&= \int_0^\infty \prod_{k\neq\ell} e^{-\left(\left(\frac{\Omega_{j\ell}}{\Omega_{jk}}\right)^{\frac{1}{1-\gamma}}\frac{\epsilon_\ell}{\lambda_k}\right)^{-\alpha}} e^{-\left(\frac{\epsilon_\ell}{\lambda_\ell}\right)^{-\alpha}}\left(\frac{\alpha}{\lambda_\ell}\right)\left(\frac{\epsilon_\ell}{\lambda_\ell}\right)^{-1-\alpha} d\epsilon_\ell \\
&= \int_0^\infty e^{-\left[\sum_{k=1}^J\left[\left(\frac{\Omega_{j\ell}}{\Omega_{jk}}\right)^{\frac{1}{1-\gamma}}\frac{1}{\lambda_k}\right]^{-\alpha}\right](\epsilon_\ell)^{-\alpha}} \frac{\alpha}{\lambda_\ell}\left(\frac{\epsilon_\ell}{\lambda_\ell}\right)^{-1-\alpha} d\epsilon_\ell \\
&= (\lambda_\ell)^\alpha \int_0^\infty e^{-\left[\sum_{k=1}^J\left[\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{1}{1-\gamma}}\lambda_k\right]^{\alpha}\right](\epsilon_\ell)^{-\alpha}} \alpha(\epsilon_\ell)^{-1-\alpha} d\epsilon_\ell
\end{aligned}$$

Change variables to $u = (\epsilon_\ell)^{-\alpha}$, and $du = -\alpha(\epsilon_\ell)^{-\alpha-1} d\epsilon_\ell$ which also requires switching the limits of integration. Thus

$$\mu_{j\ell} = -(\lambda_\ell)^\alpha \int_\infty^0 e^{-\left[\sum_{k=1}^J\left[\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{1}{1-\gamma}}\lambda_k\right]^{\alpha}\right]u} du = (\lambda_\ell)^\alpha \int_0^\infty e^{-\left[\sum_{k=1}^J\left[\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{1}{1-\gamma}}\lambda_k\right]^{\alpha}\right]u} du,$$

and therefore

$$\mu_{j\ell} = \frac{(\lambda_\ell)^\alpha}{\left[\sum_{k=1}^J\left[\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{1}{1-\gamma}}\lambda_k\right]^{\alpha}\right]}. \tag{47}$$

Expression (47) applies for all $0 \leq \gamma \neq 1$. For $\gamma < 1$, $\Omega_{j\ell} > 0$ and can be taken out as a common factor for the summation in the denominator. Doing so, using $\Omega_{j\ell}=\tau_{j\ell}^{(1-\gamma)}v^{\ell}$ and simplifying

$$\mu_{j\ell}=\frac{\left[\lambda_{\ell}\tau_{j\ell}\left(v^{\ell}\right)^{\frac{1}{1-\gamma}}\right]^{\alpha}}{\sum_{k=1}^{J}\left[\lambda_{k}\tau_{jk}\left(v^{k}\right)^{\frac{1}{1-\gamma}}\right]^{\alpha}}.$$

as claimed in the text. For $\gamma>1$, $\Omega_{j\ell}<0$, then, using absolute values, we obtain

$$\mu_{j\ell}=\frac{\left[\lambda_{\ell}\tau_{j\ell}\left(-v^{\ell}\right)^{\frac{1}{1-\gamma}}\right]^{\alpha}}{\sum_{k=1}^{J}\left[\lambda_{k}\tau_{jk}\left(-v^{k}\right)^{\frac{1}{1-\gamma}}\right]^{\alpha}}.,$$

as claimed. ■

Proof. Lemma 3: Conditional Expectation of Human Capital. By definition, the conditional expectation is

$$\begin{aligned} E\left[\epsilon_{\ell}|\Omega_{j\ell}\epsilon_{\ell}=\max_{k}\left\{\Omega_{j\ell}\left(\epsilon_{k}\right)^{1-\gamma}\right\}\right] &= \frac{\int_{0}^{\infty}\epsilon_{\ell}\prod_{k\neq\ell}\Pr\left[\epsilon_{k}<\left(\frac{\Omega_{j\ell}}{\Omega_{jk}}\right)^{\frac{1}{1-\gamma}}\epsilon_{\ell}\right]f(\epsilon_{\ell})d\epsilon_{\ell}}{\int_{0}^{\infty}\prod_{k\neq\ell}\Pr\left[\epsilon_{k}<\left(\frac{\Omega_{j\ell}}{\Omega_{jk}}\right)^{\frac{1}{1-\gamma}}\epsilon_{\ell}\right]f(\epsilon_{\ell})d\epsilon_{\ell}} \\ &= \frac{\int_{0}^{\infty}\epsilon_{\ell}\prod_{k\neq\ell}F\left[\left(\frac{\Omega_{j\ell}}{\Omega_{jk}}\right)^{\frac{1}{1-\gamma}}\epsilon_{\ell}\right]f(\epsilon_{\ell})d\epsilon_{\ell}}{\mu_{j\ell}}, \end{aligned}$$

where $\mu_{j\ell}$ is the mass of worker switching, which from Theorem 1, is given by $\mu_{j\ell}=\frac{(\lambda_{\ell})^{\alpha}}{\left[\sum_{k=1}^{J}\left[\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{1}{1-\gamma}}\lambda_{k}\right]^{\alpha}\right]}$. To solve for the numerator,

$$\begin{aligned} \int_{0}^{\infty}\epsilon_{\ell}\prod_{k\neq\ell}\Pr\left[\epsilon_{k}<\left(\frac{\Omega_{j\ell}}{\Omega_{jk}}\right)^{\frac{1}{1-\gamma}}\epsilon_{\ell}\right]f(\epsilon_{\ell})d\epsilon_{\ell} &= \int_{0}^{\infty}\epsilon_{\ell}\prod_{k\neq\ell}e^{-\left[\left(\frac{\Omega_{j\ell}}{\Omega_{jk}}\right)^{\frac{1}{1-\gamma}}\frac{\epsilon_{\ell}}{\lambda_{k}}\right]^{-\alpha}}e^{-\left(\frac{\epsilon_{\ell}}{\lambda_{\ell}}\right)^{-\alpha}}\left(\frac{\alpha}{\lambda_{\ell}}\right)\left(\frac{\epsilon_{\ell}}{\lambda_{\ell}}\right)^{-1-\alpha}d\epsilon_{\ell} \\ &= \int_{0}^{\infty}\epsilon_{\ell}e^{-\left[\sum_{k=1}^{J}\left(\left(\frac{\Omega_{j\ell}}{\Omega_{jk}}\right)^{\frac{1}{1-\gamma}}\frac{1}{\lambda_{k}}\right)^{-\alpha}\right](\epsilon_{\ell})^{-\alpha}}\frac{\alpha}{\lambda_{\ell}}\left(\frac{\epsilon_{\ell}}{\lambda_{\ell}}\right)^{-1-\alpha}d\epsilon_{\ell} \\ &= \int_{0}^{\infty}e^{-\left[\sum_{k=1}^{J}\left(\left(\frac{\Omega_{j\ell}}{\Omega_{jk}}\right)^{\frac{1}{1-\gamma}}\frac{1}{\lambda_{k}}\right)^{-\alpha}\right](\epsilon_{\ell})^{-\alpha}}\alpha\left(\lambda_{\ell}\right)^{\alpha}\left(\epsilon_{\ell}\right)^{-\alpha}d\epsilon_{\ell}. \end{aligned}$$

Now, to solve the resulting integral, change variable and define

$$z\equiv\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]\left(\epsilon_{\ell}\right)^{-\alpha}.$$

With these substitutions, $\epsilon_{\ell}=[z]^{\frac{-1}{\alpha}}\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]^{\frac{1}{\alpha}}$ and $d\epsilon_{\ell}=\frac{-1}{\alpha}[z]^{\frac{-1}{\alpha}-1}\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]^{\frac{1}{\alpha}}dz$.

With these terms plugged, the expression for the numerator becomes:

$$\begin{aligned}
&\int_{\infty}^{0} e^{-z}\alpha\left(\lambda_{\ell}\right)^{\alpha}\left([z]^{\frac{-1}{\alpha}}\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]^{\frac{1}{\alpha}}\right)^{-\alpha}\left(\frac{-1}{\alpha}\right)[z]^{\frac{-1}{\alpha}-1}\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]^{\frac{1}{\alpha}} dz \\
= & \int_{0}^{\infty} e^{-z}\left(\lambda_{\ell}\right)^{\alpha}\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]^{-1}[z]^{\frac{-1}{\alpha}}\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]^{\frac{1}{\alpha}} dz \\
= & \left(\lambda_{\ell}\right)^{\alpha}\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]^{\frac{1}{\alpha}-1}\int_{0}^{\infty} e^{-z}[z]^{\frac{-1}{\alpha}} dz \\
= & \Gamma\left(1-\frac{1}{\alpha}\right)\left(\lambda_{\ell}\right)^{\alpha}\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]^{\frac{1}{\alpha}-1} \\
= & \Gamma\left(1-\frac{1}{\alpha}\right)\lambda_{\ell}\frac{1}{\left[\left(\lambda_{\ell}\right)^{\alpha}\right]^{\frac{1-\alpha}{\alpha}}}\left[\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}\right]^{\frac{1-\alpha}{\alpha}} \\
= & \Gamma\left(1-\frac{1}{\alpha}\right)\lambda_{\ell}\left[\frac{\left(\lambda_{\ell}\right)^{\alpha}}{\sum_{k=1}^{J}\left(\frac{\Omega_{jk}}{\Omega_{j\ell}}\right)^{\frac{\alpha}{1-\gamma}}\left(\lambda_{k}\right)^{\alpha}}\right]^{\frac{\alpha-1}{\alpha}}. \\
= & \Gamma\left(1-\frac{1}{\alpha}\right)\lambda_{\ell}\left[\mu_{j\ell}\right]^{1-\frac{1}{\alpha}}.
\end{aligned}$$

Dividing by $\mu_{j\ell}$, the result attains. ■

Proof. Proposition 1: Invariant Distributions for Workers and Human Capital, Exogenous Wages. For **(a)**, notice that since $\tau_{j\ell} > 0$, then every entry of the stochastic matrix μ is positive. This implies a basic mixing condition,i.e. that for any occupation there is a positive probability to transiting to any other, i.e. for all j, ℓ, $\mu^{j\ell} > 0$. A standard result for Markov chains (e.g. Theorem 11.2. in Stokey et al. (1989)) it follows that there exists a unique invariant distribution and that it is attained from any initial distribution. Similarly for **(b)** all the entries of the matrix $\mathcal{M}$ are strictly positive; hence, from the Perron-Frobenius theorem,its largest eigenvalue has always multiplicity equal to one (i.e., not repeated), and it is real and strictly positive. Moreover, the associated eigenvector, the Perron root, has all its entries strictly positive and real. Since asymptotically, from any initial H_0, the dynamics of $H_t = H_0\mathcal{M}^t$ is dominated by the Perron root, the ratios $H_t(j)/H_t(j)$ will converge, for all j to the ratios of the Perron eigenvector. ■

Proof. Proposition 1 (Cohorts): Invariant Distributions, Exogenous Wages. The law of motion for the employment shares of workers is

$$\theta_{t+1} = \delta\theta^0 + \theta_t\mu\left(1-\delta\right),$$

which, after some obvious iterated substitutions can be writen

$$\theta_{t+1} = \delta\theta^0\sum_{\tau=0}^{t}\left[\mu\left(1-\delta\right)\right]^{\tau}.$$

The term $\sum_{\tau=0}^{t}\left[\mu\left(1-\delta\right)\right]^{\tau}$ always a converges to a finite matrix $\left[I-\mu\left(1-\delta\right)\right]^{-1}$ since the largest eigenvalue of μ is 1. Hence,

$$\theta = \delta\theta^0\left[I-\mu\left(1-\delta\right)\right]^{-1},$$

would be the invariant distribution of employment of workers across occupations.

With respect to the human capital vectors, we have

$$H_{t+1} = \delta H^0 + H_t\mathcal{M}\left(1-\delta\right),$$

i.e.:

$$H_{t+1} = \delta H^0 \sum_{\tau=0}^{t} \left[\mathcal{M}\left(1-\delta\right)\right]^{\tau}$$

If G_H, the Perron root of $\mathcal{M}$, is less than $(1-\delta)^{-1}$, then we can repeat exactly the same previous steps and the invariant distribution would be

$$H = \delta H^0 \left[I - \mathcal{M}\left(1-\delta\right)\right]^{-1}.$$

If instead $G_H\left(1-\delta\right) \geq 1$, then H_{t+1} would grow over time. If $G_H\left(1-\delta\right) > 1$, then H_t the growth is exponential. If $G_H\left(1-\delta\right) = 1$ it would be arithmetic. In any case, define the scaled and detrended human capital as

$$H_t^D = \delta H^0 \frac{\sum_{\tau=0}^{t} \left[\mathcal{M}\left(1-\delta\right)\right]^{\tau}}{\sum_{\tau=0}^{t} \left[G_H\left(1-\delta\right)\right]^{\tau}}.$$

Then, the following limits are well defined:

$$\lim_{t\to\infty} H_t^D = H^0 \times \begin{cases} \delta\left[1 - G_H\left(1-\delta\right)\right]\left[I - \mathcal{M}\left(1-\delta\right)\right]^{-1} & \text{if } G_H\left(1-\delta\right) < 1; \\ \delta \lim_{t\to\infty} \left[\frac{\sum_{\tau=0}^{t}\left[\mathcal{M}(1-\delta)\right]^{\tau}}{t}\right] & \text{if } G_H\left(1-\delta\right) = 1; \\ \delta \lim_{t\to\infty} \left[\frac{\sum_{\tau=0}^{t}\left[\mathcal{M}(1-\delta)\right]^{\tau}}{\frac{\left[G_H(1-\delta)\right]^{t}-1}{G_H(1-\delta)-1}}\right] & \text{if } G_H\left(1-\delta\right) > 1. \end{cases}$$

Note that the scaling does not change neither the eigenvalues nor the normalized eigenvectors of $\sum_{\tau=0}^{t}\left[\mathcal{M}\left(1-\delta\right)\right]^{\tau}$ for any t. In all three cases, the limit terms in the right-hand side are well defined, bounded, strictly positive and real. Hence, in each case, the Perron root is strictly positive and non-repeated. Hence, the Perron vector is also uniquely defined, real strictly positive, and once normalized, it will uniquely pin-down the asymptotic cross-section distribution of workers. ■

Proof. Proposition 2: Firm Optimization. The probability that labor is used to complete task x in occupation j at time t is,

$$\begin{aligned}
\pi_t^{H,j} &= \Pr\left(\frac{w_t^j}{z_t^{H,j}(x)} < \frac{r_t^M}{z_t^{M,j}(x)}\right) \\
&= \int_0^\infty \Pr\left(z^{M,j} < \frac{r_t^M}{w_t^j} z^{H,j}\right) f_j(z^{H,j}) dz^{H,j} \\
&= \int_0^\infty e^{-\left(\frac{r_t^M}{w_t^j}\frac{z^{H,j}}{A_t^{M,j}}\right)^{-\nu_j}} e^{-\left(\frac{z^{H,j}}{A_t^{H,j}}\right)^{-\nu_j}} \left(\frac{\nu_j}{A_t^{H,j}}\right)\left(\frac{z^{H,j}}{A_t^{H,j}}\right)^{-1-\nu_j} dz^{H,j} \\
&= \int_0^\infty e^{-\left(\frac{r_t^M}{w_t^j}\frac{z^{H,j}}{A_t^{M,j}}\right)^{-\nu_j} - \left(\frac{z^{H,j}}{A_t^{H,j}}\right)^{-\nu_j}} \left(\frac{\nu_j}{A_t^{H,j}}\right)\left(\frac{z^{H,j}}{A_t^{H,j}}\right)^{-1-\nu_j} dz^{H,j} \\
&= \nu_j \left(A_t^{H,j}\right)^{\nu_j} \int_0^\infty e^{-\left[\left(\frac{r_t^M}{w_t^j}\frac{1}{A_t^{M,j}}\right)^{-\nu_j} + \left(\frac{1}{A_t^{H,j}}\right)^{-\nu_j}\right]\left(z^{H,j}\right)^{-\nu_j}} \left(z^{H,j}\right)^{-1-\nu_j} dz^{H,j}.
\end{aligned}$$

Let $u = \left(z^{H,j}\right)^{-\nu_j}$. Then $du = -\nu_j \left(z^{H,j}\right)^{-1-\nu_j}$. With these:

$$\begin{aligned}
\pi_t^{H,j} &= -(A_t^{H,j})^{\nu_j} \int_\infty^0 e^{-\left[\left(\frac{r_t^M}{w_t^j}\frac{1}{A_t^{M,j}}\right)^{-\nu_j} + \left(\frac{1}{A_t^{H,j}}\right)^{-\nu_j}\right] u} du \\
&= (A_t^{H,j})^{\nu_j} \int_0^\infty e^{-\left[\left(\frac{r_t^M}{w_t^j}\frac{1}{A_t^{M,j}}\right)^{-\nu_j} + \left(\frac{1}{A_t^{H,j}}\right)^{-\nu_j}\right] u} du \\
&= (A_t^{H,j})^{\nu_j} \frac{1}{\left(\frac{r_t^M}{w_t^j}\frac{1}{A_t^{M,j}}\right)^{-\nu_j} + \left(\frac{1}{A_t^{H,j}}\right)^{-\nu_j}} \\
&= \frac{(w_t^j)^{-\nu_j}(A_t^{H,j})^{\nu_j}}{(r_t^M)^{-\nu_j}(A_t^{M,j})^{\nu_j} + (w_t^j)^{-\nu_j}(A_t^{H,j})^{\nu_j}},
\end{aligned}$$

as claimed in the text. Similarly, for machines, simply use the previous expression, $\pi_t^{M,j} = 1 - \pi_t^{H,j}$ and hence

$$\pi_t^{M,j} = \frac{(r_t^M)^{-\nu_j}(A_t^{M,j})^{\nu_j}}{(r_t^M)^{-\nu_j}(A_t^{M,j})^{\nu_j} + (w_t^j)^{-\nu_j}(A_t^{H,j})^{\nu_j}}.$$

Now for the unitary cost of the aggregate bundle of tasks, C_t^j, note that we can write the expression as,

$$\begin{aligned}
\left(C_t^j\right)^{(1-\eta)} &= \int_0^\infty \left(\frac{w_t^j}{z_t^{H,j}}\right)^{(1-\eta)} \Pr\left(z^{M,j} < \frac{r_t^M}{w_t^j} z^{H,j}\right) f_j(z^{H,j}) dz^{H,j} + \\
&\quad \int_0^\infty \left(\frac{r_t^M}{z_t^{M,j}}\right)^{(1-\eta)} \Pr\left(z^{H,j} < \frac{w_t^j}{r_t^M} z^{M,j}\right) f_j(z^{M,j}) dz^{M,j} \\
&= \int_0^\infty \left(\frac{w_t^j}{z_t^{H,j}}\right)^{(1-\eta)} e^{-\left(\frac{r_t^M}{w_t^j}\frac{z^{H,j}}{A_t^{M,j}}\right)^{-\nu_j}} e^{-\left(\frac{z^{H,j}}{A_t^{H,j}}\right)^{-\nu_j}} \left(\frac{\nu_j}{A_{H,t}^j}\right)\left(\frac{z^{H,j}}{A_t^{H,j}}\right)^{-1-\nu_j} dz^{H,j} + \\
&\quad \int_0^\infty \left(\frac{r_t^M}{z_t^{M,j}}\right)^{(1-\eta)} e^{-\left(\frac{w_t^j}{r_t^M}\frac{z^{M,j}}{A_t^{M,j}}\right)^{-\nu_j}} e^{-\left(\frac{z^{M,j}}{A_t^{M,j}}\right)^{-\nu_j}} \left(\frac{\nu_j}{A_t^{M,j}}\right)\left(\frac{z^{M,j}}{A_t^{M,j}}\right)^{-1-\nu_j} dz^{M,j} \\
&= (w_t^j)^{(1-\eta)} \nu_j (A_t^{H,j})^{\nu_j} \int_0^\infty e^{-\left[\left(\frac{r_t^M}{w_t^j}\frac{1}{A_t^{M,j}}\right)^{-\nu_j} + \left(\frac{1}{A_t^{H,j}}\right)^{-\nu_j}\right](z^{H,j})^{-\nu_j}} \left(z^{H,j}\right)^{-2-\nu_j+\eta} dz^{H,j} + \\
&\quad (r_t^M)^{(1-\eta)} \nu_j (A_t^{M,j})^{\nu_j} \int_0^\infty e^{-\left[\left(\frac{w_t^j}{r_t^M}\frac{1}{A_t^{H,j}}\right)^{-\nu_j} + \left(\frac{1}{A_t^{M,j}}\right)^{-\nu_j}\right](z^{M,j})^{-\nu_j}} \left(z^{M,j}\right)^{-2-\nu_j+\eta} dz^{M,j}
\end{aligned}$$

Let $u_1 = \left[\left(\frac{r_t^M}{w_t^j}\frac{1}{A_t^{M,j}}\right)^{-\nu_j} + \left(\frac{1}{A_t^{H,j}}\right)^{-\nu_j}\right](z^{H,j})^{-\nu}$, which implies that $z^{H,j} = \left[\left(\frac{r_t^M}{w_t^j}\frac{1}{A_t^{M,j}}\right)^{-\nu_j} + \left(\frac{1}{A_t^{H,j}}\right)^{-\nu_j}\right]^{1/\nu}(u_1)^{-1/\nu}$. Similarly, let $u_2 = \left[\left(\frac{w_t^j}{r_t^M}\frac{1}{A_t^{H,j}}\right)^{-\nu_j} + \left(\frac{1}{A_t^{M,j}}\right)^{-\nu_j}\right](z^{M,j})^{-\nu}$, which implies that $z^{M,j} = \left[\left(\frac{w_t^j}{r_t^M}\frac{1}{A_t^{H,j}}\right)^{-\nu_j} + \left(\frac{1}{A_t^{M,j}}\right)^{-\nu_j}\right]^{1/\nu}(u_2)^{-1/\nu}$. Then,

$$
\begin{aligned}
\left(C_t^j\right)^{(1-\eta)} &= (w_t^j)^{(1-\eta)}\,\nu_j\,(A_t^{H,j})^{\nu_j}\int_{\infty}^{0} e^{-u_1}\left(\left[\left(\frac{r_t^M}{w_t^j}\frac{1}{A_t^{M,j}}\right)^{-\nu_j}+\left(\frac{1}{A_t^{H,j}}\right)^{-\nu_j}\right]^{1/\nu_j}(u_1)^{-1/\nu_j}\right)^{-2-\nu_j+\eta} \\
&\qquad \times\left[\left(\frac{r_t^M}{w_t^j}\frac{1}{A_t^{M,j}}\right)^{-\nu_j}+\left(\frac{1}{A_t^{H,j}}\right)^{-\nu_j}\right]^{1/\nu_j}\left(\frac{-1}{\nu_j}\right)u_1^{-1-1/\nu_j}\,du_1 + \\
&\quad (r_t^M)^{(1-\eta)}\,\nu_j\,(A_t^{M,j})^{\nu_j}\int_{\infty}^{0} e^{-u_2}\left(\left[\left(\frac{w_t^j}{r_t^M}\frac{1}{A_t^{H,j}}\right)^{-\nu_j}+\left(\frac{1}{A_t^{M,j}}\right)^{-\nu_j}\right]^{1/\nu_j}\right)^{-2-\nu_j+\eta} \\
&\qquad \times\left[\left(\frac{w_t^j}{r_t^M}\frac{1}{A_t^{H,j}}\right)^{-\nu_j}+\left(\frac{1}{A_t^{M,j}}\right)^{-\nu_j}\right]^{1/\nu_j}\left(\frac{-1}{\nu_j}\right)u_2^{-1-1/\nu_j}\,du_2 \\
&= (w_t^j)^{(1-\eta)}\,(A_t^{H,j})^{\nu_j}\left[\left(\frac{r_t^M}{w_t^j}\frac{1}{A_t^{M,j}}\right)^{-\nu_j}+\left(\frac{1}{A_t^{H,j}}\right)^{-\nu_j}\right]^{\frac{-1-\nu_j+\eta}{\nu_j}}\int_0^{\infty} e^{-u_1}u_1^{\frac{1-\eta}{\nu_j}}\,du_1 + \\
&\quad (r_t^M)^{(1-\eta)}\,(A_t^{M,j})^{\nu_j}\left[\left(\frac{w_t^j}{r_t^M}\frac{1}{A_t^{H,j}}\right)^{-\nu_j}+\left(\frac{1}{A_t^{M,j}}\right)^{-\nu_j}\right]^{\frac{-1-\nu_j+\eta}{\nu_j}}\int_0^{\infty} e^{-u_2}u_2^{\frac{1-\eta}{\nu_j}}\,du_2 \\
&= (w_t^j)^{(1-\eta)}\,(A_t^{H,j})^{\nu_j}\left[\left(\frac{r_t^M}{w_t^j}\frac{1}{A_t^{M,j}}\right)^{-\nu_j}+\left(\frac{1}{A_t^{H,j}}\right)^{-\nu_j}\right]^{-1-\frac{1-\eta}{\nu_j}}\Gamma\left(1+\frac{1-\eta}{\nu_j}\right) + \\
&\quad (r_t^M)^{(1-\eta)}\,(A_t^{M,j})^{\nu_j}\left[\left(\frac{w_t^j}{r_t^M}\frac{1}{A_t^{H,j}}\right)^{-\nu_j}+\left(\frac{1}{A_t^{M,j}}\right)^{-\nu_j}\right]^{-1-\frac{1-\eta}{\nu_j}}\Gamma\left(1+\frac{1-\eta}{\nu_j}\right) \\
&= (w_t^j)^{(1-\eta)}(w_t^j)^{-\nu_j-(1-\eta)}\,(A_t^{H,j})^{\nu_j}\left[\left(\frac{r_t^M}{A_t^{M,j}}\right)^{-\nu_j}+\left(\frac{w_t^j}{A_t^{H,j}}\right)^{-\nu_j}\right]^{-1-\frac{1-\eta}{\nu_j}}\Gamma\left(1+\frac{1-\eta}{\nu_j}\right) + \\
&\quad (r_t^M)^{(1-\eta)}(r_t^M)^{-\nu_j-(1-\eta)}\,(A_t^{M,j})^{\nu_j}\left[\left(\frac{r_t^M}{A_t^{M,j}}\right)^{-\nu_j}+\left(\frac{w_t^j}{A_t^{H,j}}\right)^{-\nu_j}\right]^{-1-\frac{1-\eta}{\nu_j}}\Gamma\left(1+\frac{1-\eta}{\nu_j}\right) \\
&= (w_t^j)^{-\nu_j}\,(A_t^{H,j})^{\nu_j}\left[\left(\frac{r_t^M}{A_t^{M,j}}\right)^{-\nu_j}+\left(\frac{w_t^j}{A_t^{H,j}}\right)^{-\nu_j}\right]^{-1-\frac{1-\eta}{\nu_j}}\Gamma\left(1+\frac{1-\eta}{\nu_j}\right) + \\
&\quad (r_t^M)^{-\nu_j}\,(A_t^{M,j})^{\nu_j}\left[\left(\frac{r_t^M}{A_t^{M,j}}\right)^{-\nu_j}+\left(\frac{w_t^j}{A_t^{H,j}}\right)^{-\nu_j}\right]^{-1-\frac{1-\eta}{\nu_j}}\Gamma\left(1+\frac{1-\eta}{\nu_j}\right) \\
&= \left[\left(\frac{r_t^M}{A_t^{M,j}}\right)^{-\nu_j}+\left(\frac{w_t^j}{A_t^{H,j}}\right)^{-\nu_j}\right]\left[\left(\frac{r_t^M}{A_t^{M,j}}\right)^{-\nu_j}+\left(\frac{w_t^j}{A_t^{H,j}}\right)^{-\nu_j}\right]^{-1-\frac{1-\eta}{\nu_j}}\Gamma\left(1+\frac{1-\eta}{\nu_j}\right)
\end{aligned}
$$

Therefore,

$$
C_t^j = \Gamma\left(1+\frac{1-\eta}{\nu_j}\right)^{1/(1-\eta)}\left[(r_t^M)^{-\nu_j}(A_t^M)^{\nu_j}+(w_t^j)^{-\nu_j}(A_t^{H,j})^{\nu_j}\right]^{-\frac{1}{\nu_j}}.
$$

Finally, the formula for the unit cost of the occupation bunddle, C_t^Q, and the final good prices, P_t, follows from the well-known formulas for the unitary cost implied by a CES production and a Cobb-Douglas production functions, respectively. ■

Euler Equations of Capital Owners.

Capital owners take as given the real rental rates of structures and machines. In units of, these rental rates are denoted

$$\rho_t^K \equiv \frac{r_t^K}{P_t} \text{ and } \rho_t^{M,j} \equiv \frac{r_t^{M,j}}{P_t}.$$

Then, the Bellman Equation for capital owners is defined as:

$$V_t^K\left(K_t, M_t, B_t\right) = \max_{B_{t+1}, K_{t+1}, M_{t+1}} \left\{u\left(c_t^K\right) + \beta V_{t+1}^K\left(K_{t+1}, M_{t+1}, B_{t+1}\right)\right\}.$$

subject to

$$\begin{aligned}
c_t^K &= \rho_t^K K_t + \rho_t^M M_t + R_t B_t - B_{t+1} - I_t^K - \sum_{j=1}^{J} I_t^{M,j} \\
K_{t+1} &= \left(1-\delta^K\right) K_t + \xi_t^K I_t^K, \\
M_{t+1}^j &= \left(1-\delta^M\right) M_t^j + \xi_t^M I_t^{M,j},\ \forall j = 1, 2, \ldots J.
\end{aligned}$$

where $r_t = R_t/P_t$ is the gross, real interest rate and B_t is the net financial position (bonds) of the capital owner household. We consider **(a)** a small open economy where the rate R_t is exogenously given in world financial markets, and **(b)** a closed economy where $B_t = B_{t+1} = 0$ and the interest rate is determined by $\left\{c_t^K\right\}$. In either case, capital owners take as given $\left\{r_t,\ \rho_t^K, \rho_t^{M,j}\right\}$. The first order conditions and the envelop conditions for this problem are standard. After simplification, they lead to

$$\frac{R_{t+1}}{P_{t+1}} = \beta^{-1}\left(\frac{c_{t+1}^K}{c_t^K}\right)^{\gamma},\ \frac{r_{t+1}^K}{P_{t+1}} = \frac{\frac{R_{t+1}}{P_{t+1}} - \left(1-\delta^K\right)}{\xi_t^K}, \text{ and } \frac{r_{t+1}^M}{P_{t+1}} = \frac{\frac{R_{t+1}}{P_{t+1}} - \left(1-\delta^M\right)}{\xi_t^M},$$

which are the same expressions as in the text.

Proof. Proposition 3: Aggregation, Intertemporal Equilibrium. Using the equilibrium determination conditions (24) for any j and ℓ: Using the equilibrium determination conditions (24) (25) for any occupation j:

$$\frac{\pi_t^{H,j}}{\pi_t^{M,j}} = \frac{\left(\omega_t^j\right)^{-\nu_j}\left(A_t^{H,j}\right)^{\nu_j}}{\left(\rho_t^j\right)^{-\nu_j}\left(A_t^{M,j}\right)^{\nu_j}}.$$

Moreover, by definition of factor shares: in each of the occupations

$$\pi_t^{H,j} = \frac{\omega_t^j H_t^j}{\omega_t^j H_t^j + \rho_t^j M_t^j} \text{ and } \pi_t^{M,j} = \frac{\rho_t^j M_t^j}{\omega_t^j H_t^j + \rho_t^j M_t^j}.$$

Then:

$$\frac{\pi_t^{H,j}}{\pi_t^{M,j}} = \frac{\left(\omega_t^j\right)^{-\nu_j}\left(A_t^{H,j}\right)^{\nu_j}}{\left(\rho_t^j\right)^{-\nu_j}\left(A_t^{M,j}\right)^{\nu_j}} = \frac{\omega_t^j H_t^j}{\rho_t^j M_t^j}.$$

Using the latter equality to solve for ω_t^j/ρ_t^j

$$\frac{\omega_t^j}{\rho_t^j} = \left(\frac{M_t^j}{H_t^j}\right)^{\frac{1}{1+\nu_j}}\left(\frac{A_t^{H,j}}{A_t^{M,j}}\right)^{\frac{\nu_j}{1+\nu_j}}. \tag{48}$$

By definition, $C_t^j Q_t^j = \omega_t^j H_t^j + \rho_t^j M_t^j$. Hence, $C_t^j Q_t^j = \rho_t^j \left[\frac{\omega_t^j}{\rho_t^j} H_t^j + M_t^j\right]$, and therefore:

$$Q_t^j = \frac{\left[\frac{\omega_t^j}{\rho_t^j} H_t^j + M_t^j\right]}{C_t^j / \rho_t^j}.$$

Next notice that using the formula for the minimized cost C_t^j in Proposition 2, we obtain:

$$\frac{C_t^j}{\rho_t^j} = \Gamma\left(1 + \frac{1-\eta}{\nu_j}\right)^{\frac{1}{1-\eta}} \left[\left(\left(A_t^{M,j}\right)^{\nu_j} + \left(A_t^{H,j}\frac{\rho_t^j}{\omega_t^j}\right)^{\nu_j}\right)\right]^{\frac{-1}{\nu_j}}.$$

Using the last equation into the last formula for Q_t^j

$$Q_t^j = \frac{\left[\frac{\omega_t^j}{\rho_t^j} H_t^j + M_t^j\right]}{\Gamma\left(1 + \frac{1-\eta}{\nu_j}\right)^{\frac{1}{1-\eta}} \left[\left(\left(A_t^{M,j}\right)^{\nu_j} + \left(A_t^{H,j}\frac{\rho_t^j}{\omega_t^j}\right)^{\nu_j}\right)\right]^{\frac{-1}{\nu_j}}}.$$

With this, we can rearrange and then substitute the formula (48) for ω_t^j / ρ_t^j

$$\begin{aligned} Q_t^j &= \Gamma\left(1 + \frac{1-\eta}{\nu_j}\right)^{\frac{1}{\eta-1}} \frac{\left[\left(\frac{M_t^j}{H_t^j}\right)^{\frac{1}{1+\nu_j}} \left(\frac{A_t^{H,j}}{A_t^{M,j}}\right)^{\frac{\nu_j}{1+\nu_j}} H_t^j + M_t^j\right]}{\left[\left(A_t^{M,j}\right)^{\nu_j} + \left(\frac{A_t^{H,j}}{\left(\frac{M_t^j}{H_t^j}\right)^{\frac{1}{1+\nu_j}} \left(\frac{A_t^{H,j}}{A_t^{M,j}}\right)^{\frac{\nu_j}{1+\nu_j}}}\right)^{\nu_j}\right]^{\frac{-1}{\nu_j}}}, \\ &= \Gamma\left(1 + \frac{1-\eta}{\nu_j}\right)^{\frac{1}{\eta-1}} \left[\left(A_t^{M,j} M_t^j\right)^{\frac{\nu_j}{1+\nu_t^j}} + \left(A_t^{H,j} H_t^j\right)^{\frac{\nu_j}{1+\nu_t^j}}\right]^{\frac{1+\nu_j}{\nu_t^j}}. \end{aligned}$$

The last formula obtains after regrouping and simplifying the previous equation, and coincides with the one in the text. The formulae for Q_t and for Y_t result from simply using the CES and the Cobb-Douglass aggregators, respectively.

For part (b), the formulae for the rental rates and unitary wages, obtain from direct derivation of Y_t with respect to the respective factor of production and then simplifying. ■

Proof. Theorem 2: Existence of an equilibrium BGP.

To establish existence, we use the equilibrium conditions of the model to construct a continuous mapping f from a closed set to itself. We will work with a closed set of machine-to-worker ratios, $x_j \equiv \frac{M^j}{H^j}$ for $j = 1, \ldots J$. We explain why a fixed point of such a mapping induces a solution for all the other relevant variables for an equilibrium BGP.

Recall that total output is given by

$$Y_t = (K_t)^{\varphi} \left[\sum_{j=1}^{J} \Gamma_j \left[\left(A_t^{M,j} M_t^j\right)^{\frac{\nu_j}{1+\nu_j}} + \left(A_t^{H,j} H_t^j\right)^{\frac{\nu_j}{1+\nu_j}}\right]^{\frac{(1+\nu_j)(\rho-1)}{\nu_j \rho}}\right]^{\frac{(1-\varphi)\rho}{\rho-1}}, \tag{49}$$

where we have kept the notation $\Gamma_j \equiv \Gamma\left(1 + \frac{1-\eta}{\nu_j}\right)^{\frac{\rho-1}{\rho(\eta-1)}}$ to shorten the expressions.

We consider a BGP in which

$$A_t^{H,j} = a^{H,j}\left(G_A\right)^t, \text{ and } A_t^{M,j} = a^{M,j},$$

where $a^{H,j} > 0$, $a^{M,j} > 0$, and $G_A \geq 1$. If $G_H < (1-\delta)^{-1}$, the economy settles in a steady state vector of human capitals H^j, total output will grow over time at the rate $G_Y = G_A$. If instead $G_H > (1-\delta)^{-1}$, then there would be sustained growth in the vector of aggregate human capital, which, in a BGP would be given by $H_t^j = H^j \left[G_H (1-\delta)\right]^t$. In this case, aggregate ouput Y_t and aggregate structures and machines K_t, M_t^j would all grow at the rate and grow at the rate $G_Y \equiv G_H (1-\delta) G_A$. With this definition of G_Y, total output would follow:

$$Y_t = [K_t]^{\varphi} \left[\sum_{j=1}^{J} \Gamma_j \left\{ \left[M_t^j \cdot a^{M,j} \right]^{\frac{\nu_j}{1+\nu_j}} + \left[H^j \cdot a^{H,j} \left[G_Y\right]^t \right]^{\frac{\nu_j}{1+\nu_j}} \right\}^{\frac{(1+\nu_j)(\rho-1)}{\nu_j \rho}} \right]^{\frac{(1-\varphi)\rho}{\rho-1}}. \tag{50}$$

Then, in either form of BGP, the rental rates of structures, ρ^K, machines, $\rho^{M,j}$, and of unitary skills prices w_t^j, will be given by

$$\begin{aligned}
\rho^K &= \varphi \frac{Y_t}{K_t}, \\
\rho^{M,j} &= (1-\varphi) [K_t]^{\varphi} \Theta_t \Gamma_j \left\{ \left[\frac{M_t^j a^{M,j}}{(G_Y)^t} \right]^{\frac{\nu_j}{1+\nu_j}} + \left[H^j a^{H,j} \right]^{\frac{\nu_j}{1+\nu_j}} \right\}^{\frac{(1+\nu_j)(\rho-1)}{\nu_j \rho}-1} \left[a^{M,j} \right]^{\frac{\nu_j}{1+\nu_j}} \left[(G_Y)^t \right]^{1-\varphi-\frac{\nu_j}{1+\nu_j}} \left[M_t^j \right]^{\frac{\nu_j}{1+\nu_j}-1}, \\
w_t^j &= (1-\varphi) [K_t]^{\varphi} \Theta_t \Gamma_j \left\{ \left[\frac{M_t^j a^{M,j}}{(G_Y)^t} \right]^{\frac{\nu_j}{1+\nu_j}} + \left[H^j a^{H,j} \right]^{\frac{\nu_j}{1+\nu_j}} \right\}^{\frac{(1+\nu_j)(\rho-1)}{\nu_j \rho}-1} \left[a^{H,j} \right]^{\frac{\nu_j}{1+\nu_j}} \left[(G_A)^t \right]^{1-\varphi} \left[H^j \right]^{\frac{\nu_j}{1+\nu_j}-1},
\end{aligned}$$

where $\Theta_t = \left\{ \sum_{k=1}^{J} \Gamma_k \left(\left[M_t^k a^{M,k} \left(\frac{1}{G_Y} \right)^t \right]^{\frac{\nu_k}{1+\nu_k}} + \left[H^k a^{H,k} \right]^{\frac{\nu_k}{1+\nu_k}} \right)^{\frac{(1+\nu_k)(\rho-1)}{\nu_k \rho}} \right\}^{\frac{(1-\varphi)\rho}{\rho-1}-1}$.

The first equation indicates that the output-to-structures ratio must be constant along the BGP. Hence, K_t must grow at the rate of growth of Y_t. The second equation indicates that for constant rental rates for machines, then, the stock of machines M_t^j must grow at the rate G_Y. In this way, the term $M_t^j \cdot a^{M,j} (G_Y)^{-t}$ is constant over time. Let $M_t^j = m^j (G_Y)^t$ and $K_t = k (G_Y)^t$, then, after simplifying:

$$\rho^{M,j} = (1-\varphi) [k]^{\varphi} \Phi_t \Gamma_j \left\{ \left[m^j a^{M,j} \right]^{\frac{\nu_j}{1+\nu_j}} + \left[H^j a^{H,j} \right]^{\frac{\nu_j}{1+\nu_j}} \right\}^{\frac{(1+\nu_j)(\rho-1)}{\nu_j \rho}-1} \left[a^{M,j} \right]^{\frac{\nu_j}{1+\nu_j}} \left[m^j \right]^{\frac{\nu_j}{1+\nu_j}-1}, \text{ and} \tag{51}$$

$$w_t^j = (1-\varphi) [k]^{\varphi} \Phi_t \Gamma_j \left\{ \left[m^j a^{M,j} \right]^{\frac{\nu_j}{1+\nu_j}} + \left[H^j a^{H,j} \right]^{\frac{\nu_j}{1+\nu_j}} \right\}^{\frac{(1+\nu_j)(\rho-1)}{\nu_j \rho}-1} \left[a^{H,j} \right]^{\frac{\nu_j}{1+\nu_j}} \left[H^j \right]^{\frac{\nu_j}{1+\nu_j}-1} (G_A)^t. \tag{52}$$

With exogenously given time invariant and strictly positive productivities $\left\{ a^{M,\,j}, a^{H,\,,j} \right\}_{j=1}^{J}$, we now use equations (51) and (52) to define the continuous mapping f from a closed set of machine-to-worker ratios, $x_j = m^j / H^j$, $j = 1, ...J$, to itself, i.e.: $f : \Delta^J \to \Delta^J$, where $\Delta^J \equiv \left\{ x \in \mathbb{R}_+^J : \sum_{j=1}^{J} x_j = 1 \right\}$ is the J-dimensional simplex. To start, fix any $x \in \Delta^J$ such that $x_j > 0$ for all $j = 1, 2, ...J$, and re-write (52) by (51), in terms of x:

$$\rho^{M,j} = (1-\varphi) [k]^{\varphi} \Phi_t \Gamma_j \left[a^{M,j} \right]^{\frac{\nu_j}{1+\nu_j}} \left[H^j \right]^{-\frac{1}{\rho}} \left\{ \left(\left[x_j a^{M,j} \right]^{\frac{\nu_j}{1+\nu_j}} + \left[a^{H,j} \right]^{\frac{\nu_j}{1+\nu_j}} \right) \right\}^{\frac{(1+\nu_j)(\rho-1)}{\nu_j \rho}-1} \left[x_j \right]^{\frac{\nu_j}{1+\nu_j}-1}, \tag{53}$$

$$w_t^j = (1-\varphi) [k]^{\varphi} \Phi_t \Gamma_j \left[a^{H,j} \right]^{\frac{\nu_j}{1+\nu_j}} \left[H^j \right]^{\frac{-1}{\rho}} \left\{ \left(\left[x_j a^{M,j} \right]^{\frac{\nu_j}{1+\nu_j}} + \left[a^{H,j} \right]^{\frac{\nu_j}{1+\nu_j}} \right) \right\}^{\frac{(1+\nu_j)(\rho-1)}{\nu_j \rho}-1} (G_A)^t. \tag{54}$$

For any j, divide (54) by (53) and obtain:

$$w_t^j = \rho^{M,j} \left(\frac{a^{H,j}}{a^{M,j}} \right)^{\frac{\nu_j}{1+\nu_j}} (x_j)^{\frac{1}{1+\nu_j}} (G_A)^t .$$

That is, unitary wages in occupation j depend on the rental rate of machines $\rho^{M,j}$, the relative productivities of those two factors, and the machine-to-workers ratio x_j specific to that occupation and a common growth term. But since the rental rates of machines across all occupations are equal, for any $k \neq j$, then

$$\frac{\omega^j}{\omega^k}(x) \equiv \frac{w_t^j}{w_t^k} = \frac{\left(\frac{a^{H,j}}{a^{M,j}} \right)^{\frac{\nu_j}{1+\nu_j}} (x_j)^{\frac{1}{1+\nu_j}}}{\left(\frac{a^{H,k}}{a^{M,k}} \right)^{\frac{\nu_k}{1+\nu_k}} (x_k)^{\frac{1}{1+\nu_k}}},$$

which is positive and well defined since $x_j > 0$ and $x_k > 0$. Then, $\frac{\omega^j}{\omega^k}(x)$, the relative wage of occupation j relative to occupation k is constant over time, and a continuous function of x. Recall that, the relative wages w^j/w^k uniquely determine the value of the relative supply H^j/H^k because of the homogeneity of the dynamic programming problem of workers. Specifically, following Proposition 2, let $H^0(\omega)$ denote the vector of human capitals from new workers, implied by their initial occupation choices, given $\omega(x)$. Then, define:

$$H^P(x) \equiv H^0[\omega(x)] \times \begin{cases} [1 - G_H[\omega(x)](1-\delta)][I - \mathcal{M}[\omega(x)](1-\delta)]^{-1} & \text{if } G_H[\omega(x)](1-\delta) < 1, \\ \lim_{t\to\infty} \left[\frac{\sum_{\tau=0}^{t} [\mathcal{M}[\omega(x)](1-\delta)]^\tau}{t} \right] & \text{if } G_H[\omega(x)](1-\delta) = 1, \\ \lim_{t\to\infty} \left[\frac{\sum_{\tau=0}^{t} [\mathcal{M}[\omega(x)](1-\delta)]^\tau}{\frac{[G_H[\omega(x)](1-\delta)]^t - 1}{G_H[\omega(x)](1-\delta) - 1}} \right] & \text{if } G_H[\omega(x)](1-\delta) > 1, \end{cases}$$

where $G_H(\omega(x))$ is the Perron root of $\mathcal{M}(\omega(x))$. For each j, define the function $H_j^P(x)$ as the $j-th$ coordinate of $H^P(x)$, which, up to a scale, defines the human capital in occupation j as a continuous function of x. In sum, with the equilibrium x we can solve for the aggregate supply, using the equilibrium occupation choices and human capital accumulation of workers along a BGP.

Similarly, given x, we now show how the equalization in the rental rates of machines define a continuous mapping for the relative investment in machines across occupations. For any two $j \neq k$, equate their rental rate of machines (53), after cancelling common terms and rearranging and obtain:

$$\frac{x_j}{x_k} = \frac{\Gamma_j \left[a^{M,j}\right]^{\frac{\nu_j}{1+\nu_j}} \left[H^j\right]^{-\frac{1}{\rho}} \left\{ \left(\left[x_j a^{M,j}\right]^{\frac{\nu_j}{1+\nu_j}} + \left[a^{H,j}\right]^{\frac{\nu_j}{1+\nu_j}} \right) \right\}^{\frac{(1+\nu_j)(\rho-1)}{\nu_j \rho} - 1} [x_j]^{\frac{\nu_j}{1+\nu_j}}}{\Gamma_k \left[a^{M,k}\right]^{\frac{\nu_k}{1+\nu_k}} \left[H^k\right]^{-\frac{1}{\rho}} \left\{ \left(\left[x_k a^{M,k}\right]^{\frac{\nu_k}{1+\nu_k}} + \left[a^{H,k}\right]^{\frac{\nu_k}{1+\nu_k}} \right) \right\}^{\frac{(1+\nu_k)(\rho-1)}{\nu_k \rho} - 1} [x_k]^{\frac{\nu_k}{1+\nu_k}}}.$$

We use this condition to define the function of relative investments in machines for occupations j relative to machines in occupation k that would equate their rental rates in a BGP:

$$M_{j,k}(x) \equiv \frac{\Gamma_j}{\Gamma_k} \left[\frac{H_j^P(x)}{H_k^P(x)} \right]^{-\frac{1}{\rho}} \frac{\left(\left[x_j a^{M,j}\right]^{\frac{\nu_j}{1+\nu_j}} + \left[a^{H,j}\right]^{\frac{\nu_j}{1+\nu_j}} \right)^{\frac{(1+\nu_j)(\rho-1)}{\nu_j \rho} - 1} \left[a^{M,j} x_j\right]^{\frac{\nu_j}{1+\nu_j}}}{\left(\left[x_k a^{M,k}\right]^{\frac{\nu_k}{1+\nu_k}} + \left[a^{H,k}\right]^{\frac{\nu_k}{1+\nu_k}} \right)^{\frac{(1+\nu_k)(\rho-1)}{\nu_k \rho} - 1} \left[a^{M,k} x_k\right]^{\frac{\nu_k}{1+\nu_k}}}.$$

For any fixed $k \in \{1, ...J\}$, define $F_{j,k}(x) \equiv \frac{M_{j,k}(x)}{H_j^P(x)}$ for all j. $F_{j,k}(x)$ is positive for all j and continuous on x.

Scaling it as

$$f(x) \equiv \frac{F_{j,k}(x)}{\sum_{\ell=1}^{J} F_{\ell,k}(x)},$$

implies $f(x) \in \Delta^J$. It also makes the function independent of the chosen denominator k and thus defines a continuous mapping $f : \Delta^J \to \Delta^J$. From Brower's fixed point theorem (e.g. Mascolell et al. (1995), page 962) there exists at least one fixed point $x \in \Delta^J$, i.e. $x = f(x)$.

Having determined a fixed point x, then the level of the productivities $\left\{a^{M,j}, a^{H,j}\right\}_{j=1}^{J}$ determines the relative level of Q/K, and given ρ^k, the determination of the value of the constant $k = K_t/Y_t$ is standard and straightforward. ■

B Dynamic Hat Algebra

The following equations describe the set of equilibrium conditions of the model. For simplicity we take the case of $\gamma > 1$ and a small open economy with cohorts (stochastic lifetimes), which we use in our quantitative application.

$$v_t^j = \frac{(w_t^j/P_t)^{1-\gamma}}{1-\gamma} - \beta\,\Gamma\left(1-\frac{1-\gamma}{\alpha}\right)\left[\sum_{\ell=1}^{J}(-v_{t+1}^{\ell})^{\frac{\alpha}{1-\gamma}}\ (\tau_{j\ell}\lambda_\ell)^{\alpha}\right]^{\frac{1-\gamma}{\alpha}} \tag{55}$$

$$\mu_t^{j\ell} = \frac{\left[\lambda_\ell\tau_{j\ell}\left(-v_{t+1}^{\ell}\right)^{\frac{1}{1-\gamma}}\right]^{\alpha}}{\sum_{k=1}^{J}\left[\lambda_k\tau_{jk}\left(-v_{t+1}^{k}\right)^{\frac{1}{1-\gamma}}\right]^{\alpha}} \tag{56}$$

$$H_{t+1}^{\ell} = (1-\delta)\sum_{j=1}^{J}\Gamma\left(1-\frac{1}{\alpha}\right)\tau_{j\ell}\,\lambda_\ell\left(\mu_t^{j\ell}\right)^{1-\frac{1}{\alpha}}\ H_t^j + \delta H_{t+1}^{0,\ell} \tag{57}$$

$$\theta_{t+1}^{\ell} = (1-\delta)\sum_{j=1}^{J}\mu_t^{j\ell}\,\theta_t^j + \delta\theta_{t+1}^{0,\ell} \tag{58}$$

$$\theta_{t+1}^{0,j} = \frac{\left[\lambda_j\tau^{0,j}\left(-v_{t+1}^{j}\right)^{\frac{1}{1-\gamma}}\right]^{\alpha}}{\sum\limits_{k=1}^{J}\left[\lambda_k\tau^{0,k}\left(-v_{t+1}^{k}\right)^{\frac{1}{1-\gamma}}\right]^{\alpha}}, \tag{59}$$

$$H_{t+1}^{0,j} = \Gamma\left(1-\frac{1}{\alpha}\right)\tau^{0,j}\,\lambda_j\left[\theta_{t+1}^{0,j}\right]^{1-\frac{1}{\alpha}} \tag{60}$$

$$\frac{w_t^j}{P_t} = \Gamma\left(1+\frac{1-\eta}{\nu_j}\right)^{\frac{-\nu_j}{(1+\nu_j)(1-\eta)}}(1-\varphi)^{\frac{\rho}{1+\nu_j}}\left(A_t^{jH}\right)^{\frac{\nu_j}{(1+\nu_j)}}\left(H_t^j\right)^{-\frac{1}{(1+\nu_j)}}K_t^{\frac{\varphi(\rho-1)}{(1-\varphi)(1+\nu_j)}}\left(\frac{C_t^j}{P_t}\right)^{\frac{1+\nu_j-\rho}{1+\nu_j}}(Y_t)^{\frac{1-\rho\varphi}{(1+\nu_j)(1-\varphi)}} \tag{61}$$

$$\frac{r_t^M}{P_t} = \Gamma\left(1+\frac{1-\eta}{\nu_j}\right)^{\frac{-\nu_j}{(1+\nu_j)(1-\eta)}}\left(A_t^{jM}\right)^{\frac{\nu_j}{(1+\nu_j)}}\left(M_t^j\right)^{-\frac{1}{(1+\nu_j)}}\left(\frac{C_t^j}{P_t}\right)^{1-\frac{\rho}{1+\nu_j}}((1-\varphi)Y_t)^{\frac{\rho}{1+\nu_j}}\left(\frac{Y_t^{1/(1-\varphi)}}{K_t^{\varphi/(1-\varphi)}}\right)^{\frac{1-\rho}{1+\nu_j}} \tag{62}$$

$$\frac{r_t^K}{P_t} = \varphi\frac{Y_t}{K_t} \tag{63}$$

$$\frac{C_t^j}{P_t} = \Gamma\left(1+\frac{1-\eta}{\nu_j}\right)^{\frac{1}{1-\eta}}\left[\left(\frac{A_t^{M,j}}{r_t^M/P_t}\right)^{\nu_j} + \left(\frac{A_t^{H,j}}{w_t^j/P_t}\right)^{\nu_j}\right]^{\frac{-1}{\nu_j}} \tag{64}$$

$$Y_t = (K_t)^{\varphi}\left[\sum_{j=1}^{J}\Gamma\left(1+\frac{1-\eta}{\nu_j}\right)^{\frac{\rho-1}{\rho(\eta-1)}}\left[\left(A_t^{jM}M_t^j\right)^{\frac{\nu_j}{1+\nu_j}} + \left(A_t^{jH}H_t^j\right)^{\frac{\nu_j}{1+\nu_j}}\right]^{\frac{(1+\nu_j)(\rho-1)}{\nu_j\,\rho}}\right]^{\frac{(1-\varphi)\rho}{\rho-1}} \tag{65}$$

$$\frac{R_{t+1}}{P_{t+1}} = \beta^{-1}\left(\frac{c_{t+1}^K}{c_t^K}\right)^{\gamma} \tag{66}$$

$$\frac{r_{t+1}^K}{P_{t+1}} = \frac{R_{t+1}}{P_{t+1}} - (1-\delta^K) \tag{67}$$

$$\frac{r_{t+1}^M}{P_{t+1}} = \left(\xi_{t+1}^M\right)^{-1}\left[\frac{R_{t+1}}{P_t} - (1-\delta^M)\right] \tag{68}$$

$$\tag{69}$$

We now operate over these equations to achieve a more tractable set of equilibrium conditions. Take the ratio of v_{t+1}^j and v_t^j using (55).

$$\begin{aligned}
\frac{v_{t+1}^j}{v_t^j} &= \left(\frac{w_{t+1}^j/P_{t+1}}{w_t^j/P_t}\right)^{1-\gamma} \frac{(w_t^j/P_t)^{1-\gamma}}{(1-\gamma)v_t^j} - \frac{\beta\,\Gamma\left(1-\frac{1-\gamma}{\alpha}\right)}{v_t^j}\left[\sum_{\ell=1}^{J}(-v_{t+1}^\ell)^{\frac{\alpha}{1-\gamma}}\;(\tau_{j\ell}\lambda_\ell)^\alpha\right]^{\frac{1-\gamma}{\alpha}} \frac{\left[\sum_{\ell=1}^{J}(-v_{t+2}^\ell)^{\frac{\alpha}{1-\gamma}}\;(\tau_{j\ell}\lambda_\ell)^\alpha\right]^{\frac{1-\gamma}{\alpha}}}{\left[\sum_{\ell=1}^{J}(-v_{t+1}^\ell)^{\frac{\alpha}{1-\gamma}}\;(\tau_{j\ell}\lambda_\ell)^\alpha\right]^{\frac{1-\gamma}{\alpha}}} \\
&= \left(\frac{w_{t+1}^j/P_{t+1}}{w_t^j/P_t}\right)^{1-\gamma}\Phi_t^j + \left(1-\Phi_t^j\right)\left[\sum_{\ell=1}^{J}\frac{(-v_{t+1}^\ell)^{\frac{\alpha}{1-\gamma}}\;(\tau_{j\ell}\lambda_\ell)^\alpha}{\left[\sum_{k=1}^{J}(-v_{t+1}^k)^{\frac{\alpha}{1-\gamma}}\;(\tau_{jk}\lambda_k)^\alpha\right]}\frac{(-v_{t+2}^\ell)^{\frac{\alpha}{1-\gamma}}}{(-v_{t+1}^k)^{\frac{\alpha}{1-\gamma}}}\right]^{\frac{1-\gamma}{\alpha}} \\
&= \left(\frac{w_{t+1}^j/P_{t+1}}{w_t^j/P_t}\right)^{1-\gamma}\Phi_t^j + \left(1-\Phi_t^j\right)\left[\sum_{\ell=1}^{J}\mu_t^{j\ell}\left(\frac{v_{t+2}^\ell}{v_{t+1}^\ell}\right)^{\frac{\alpha}{1-\gamma}}\right]^{\frac{1-\gamma}{\alpha}} \\
\hat{v}_{t+1}^j &= \Phi_t^j\;\left(\hat{w}_{t+1}^j/\hat{P}_{t+1}\right)^{1-\gamma} + \left(1-\Phi_t^j\right)\left[\sum_{\ell=1}^{J}\mu_t^{j\ell}\left(\hat{v}_{t+2}^\ell\right)^{\frac{\alpha}{1-\gamma}}\right]^{\frac{1-\gamma}{\alpha}},
\end{aligned}$$

where, $\Phi_t^j = \frac{(w_t^j/P_t)^{1-\gamma}}{(1-\gamma)v_t^j}$.

Similarly, take the ratio between $\mu_t^{j\ell}$ and $\mu_{t-1}^{j\ell}$ using (56),

$$\begin{aligned}
\frac{\mu_t^{j\ell}}{\mu_{t-1}^{j\ell}} &= \frac{\left(\frac{v_{t+1}^\ell}{v_t^\ell}\right)^{\frac{\alpha}{1-\gamma}}}{\sum_{k=1}^{J}\frac{\lambda_k^\alpha\tau_{jk}^\alpha\left(-v_{t+1}^k\right)^{\frac{\alpha}{1-\gamma}}}{\sum_{m=1}^{J}\lambda_m^\alpha\tau_{jm}^\alpha(-v_t^m)^{\frac{\alpha}{1-\gamma}}}} \\
&= \frac{\left(\frac{v_{t+1}^\ell}{v_t^\ell}\right)^{\frac{\alpha}{1-\gamma}}}{\sum_{k=1}^{J}\mu_t^{j\ell}\left(\frac{v_{t+1}^\ell}{v_t^\ell}\right)^{\frac{\alpha}{1-\gamma}}} \\
&= \frac{\left(\hat{v}_{t+1}^\ell\right)^{\frac{\alpha}{1-\gamma}}}{\sum_{k=1}^{J}\mu_{t-1}^{j\ell}\left(\hat{v}_{t+1}^\ell\right)^{\frac{\alpha}{1-\gamma}}}.
\end{aligned}$$

Call $\mathcal{M}_t^{j\ell} = \Gamma\left(1-\frac{1}{\alpha}\right)\,\lambda^\ell\,\tau^{j\ell}(\mu_t^{j\ell})^{1-\frac{1}{\alpha}}$, then we can write equation (57),

$$H_{t+1}^\ell = (1-\delta)\sum_{j=1}^{J}\left(\frac{\mu_t^{j\ell}}{\mu_{t-1}^{j\ell}}\right)^{1-\frac{1}{\alpha}}\;\mathcal{M}_{t-1}^{j\ell}\;H_t^j + \delta H_{t+1}^{0,\ell}.$$

And for the entering cohort, take the ratio of $\theta_{t+1}^{0,j}$ and $\theta_t^{0,j}$,

$$\frac{\theta_{t+1}^{0,j}}{\theta_t^{0,j}} = \frac{\left(\frac{v_{t+1}^j}{v_t^j}\right)^{\frac{\alpha}{1-\gamma}}}{\sum_{k=1}^{J}\theta_t^{0,k}\left(\frac{v_{t+1}^k}{v_t^k}\right)^{\frac{\alpha}{1-\gamma}}} \tag{70}$$

$$= \frac{\left(\hat{v}_{t+1}^j\right)^{\frac{\alpha}{1-\gamma}}}{\sum_{k=1}^{J}\theta_t^{0,k}\left(\hat{v}_{t+1}^k\right)^{\frac{\alpha}{1-\gamma}}} \tag{71}$$

For output, we start with the expression (65).

$$Y_t = (K_t)^\varphi\left(Q_t^{1-\varphi}\right)$$

where

$$Q_t = \left[\sum_{j=1}^{J} \Gamma \left(1 + \frac{1-\eta}{\nu_j} \right)^{\frac{\rho-1}{\rho(\eta-1)}} \left[\left(A_t^{jM} M_t^j \right)^{\frac{\nu_j}{1+\nu_j}} + \left(A_t^{jH} H_t^j \right)^{\frac{\nu_j}{1+\nu_j}} \right]^{\frac{(1+\nu_j)(\rho-1)}{\nu_j \, \rho}} \right]^{\frac{\rho}{\rho-1}}$$

and the hat algebra is

$$\hat{Y}_{t+1} = \left(\hat{K}_{t+1} \right)^{\varphi} \left(\hat{Q}_{t+1}^{1-\varphi} \right).$$

Define Q_t^j as,

$$Q_t^j = \Gamma \left(1 + \frac{1-\eta}{\nu_j} \right)^{\frac{1}{\eta-1}} \left[\left(A_t^{jM} M_t^j \right)^{\frac{\nu_j}{1+\nu_j}} + \left(A_t^{jH} H_t^j \right)^{\frac{\nu_j}{1+\nu_j}} \right]^{\frac{1+\nu_j}{\nu_j}}$$

and note that, by definition of the CES aggregator,

$$Q_t = \left[\sum_{j=1}^{J} \left(Q_t^j \right)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}} \tag{72}$$

then, the hat algebra is

$$\hat{Q}_{t+1} = \left[\sum_{j=1}^{J} \left(\hat{Q}_{t+1}^j \right)^{\frac{\rho-1}{\rho}} \left(\frac{Q_t^j}{Q_t} \right)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}}$$

note that by (72), $\left[\sum_{j=1}^{J} \left(\frac{Q_t^j}{Q_t} \right)^{\frac{\rho-1}{\rho}} \right] = 1$ and each of these terms in the sum is a share of the total (by properties of the CES function, is the share of total expenditures in Q^j, i.e. $C_t^j Q_t^j$, out of the total value of Q_t, i.e. $C_t^Q Q_t$. For now, define

$$\varrho_t^j = \left(\frac{Q_t^j}{Q_t} \right)^{\frac{\rho-1}{\rho}}.$$

The hat algebra for Q_t^j is,

$$\hat{Q}_{t+1}^j = \Gamma \left(1 + \frac{1-\eta}{\nu_j} \right)^{\frac{1}{\eta-1}} \left[\left(\hat{A}_{t+1}^{jM} \hat{M}_{t+1}^j \right)^{\frac{\nu_j}{1+\nu_j}} \left(\frac{A_t^{jM} M_t^j}{Q_t^j} \right)^{\frac{\nu_j}{1+\nu_j}} + \left(\hat{A}_{t+1}^{jH} \hat{H}_{t+1}^j \right)^{\frac{\nu_j}{1+\nu_j}} \left(\frac{A_t^{jH} H_t^j}{Q_t^j} \right)^{\frac{\nu_j}{1+\nu_j}} \right]^{\frac{1+\nu_j}{\nu_j}}.$$

Using the following identity, which can be easily derived using the definition of π and the equilibrium shares,

$$\pi_t^{jM} = \left(\frac{\Gamma \left(1 + \frac{1-\eta}{\nu_j} \right)^{\frac{1}{\eta-1}} A_t^{jM} M_t^j}{Q_t^j} \right)^{\frac{\nu_j}{1+\nu_j}} \tag{73}$$

and similarly for π_t^{jH}, we can rewrite the previous equation as,

$$\hat{Q}_{t+1}^j = \left[\left(\hat{A}_{t+1}^{jM} \hat{M}_{t+1}^j \right)^{\frac{\nu_j}{1+\nu_j}} \pi_t^{jM} + \left(\hat{A}_{t+1}^{jH} \hat{H}_{t+1}^j \right)^{\frac{\nu_j}{1+\nu_j}} \pi_t^{jH} \right]^{\frac{1+\nu_j}{\nu_j}}$$

Combining all the previous results we have,

$$\hat{Y}_{t+1} = \left(\hat{K}_{t+1}\right)^{\varphi} \left[\sum_{j=1}^{J} \varrho_t^j \left[\left(\hat{A}_{t+1}^{jM}\hat{M}_{t+1}^j\right)^{\frac{\nu_j}{1+\nu_j}} \pi_t^{jM} + \left(\hat{A}_{t+1}^{jH}\hat{H}_{t+1}^j\right)^{\frac{\nu_j}{1+\nu_j}} \pi_t^{jH}\right]^{\frac{(1+\nu_j)(\rho-1)}{\nu_j \rho}}\right]^{\frac{(1-\varphi)\rho}{\rho-1}}.$$

And the law of motion for ϱ_t^j is,

$$\varrho_{t+1}^j = \frac{\varrho_t^j \left[\left(\hat{A}_{t+1}^{jM}\hat{M}_{t+1}^j\right)^{\frac{\nu_j}{1+\nu_j}} \pi_t^{jM} + \left(\hat{A}_{t+1}^{jH}\hat{H}_{t+1}^j\right)^{\frac{\nu_j}{1+\nu_j}} \pi_t^{jH}\right]^{\frac{(1+\nu_j)(\rho-1)}{\nu_j \rho}}}{\sum_{j=1}^{J} \varrho_t^j \left[\left(\hat{A}_{t+1}^{jM}\hat{M}_{t+1}^j\right)^{\frac{\nu_j}{1+\nu_j}} \pi_t^{jM} + \left(\hat{A}_{t+1}^{jH}\hat{H}_{t+1}^j\right)^{\frac{\nu_j}{1+\nu_j}} \pi_t^{jH}\right]^{\frac{(1+\nu_j)(\rho-1)}{\nu_j \rho}}}.$$

We now derive the hat algebra for wages, rental rates and unit costs. The expressions for real wages and rental rates use (73) and

$$\frac{r_t^M}{P_t} M_t^j = \pi_t^{jM} \frac{C_t^j}{P_t} Q_t^j; \qquad \frac{w_t^j}{P_t} H_t^j = \pi_t^{jH} \frac{C_t^j}{P_t} Q_t^j.$$

After some algebra, it follows that

$$\frac{r_t^M}{P_t} = \Gamma \left(1 + \frac{1-\eta}{\nu_j}\right)^{\frac{-\nu_j}{(1+\nu_j)(1-\eta)}} \left(A_t^{jM}\right)^{\frac{\nu_j}{(1+\nu_j)}} \left(M_t^j\right)^{-\frac{1}{(1+\nu_j)}} \frac{C_t^j}{P_t} \left(Q_t^j\right)^{\frac{1}{1+\nu_j}}$$

using the optimal CES demand, $Q_t^j = \left(\frac{C_t^j/P_t}{C_t^Q/P_t}\right)^{-\rho} Q_t$, we have

$$\frac{r_t^M}{P_t} = \Gamma \left(1 + \frac{1-\eta}{\nu_j}\right)^{\frac{-\nu_j}{(1+\nu_j)(1-\eta)}} \left(A_t^{jM}\right)^{\frac{\nu_j}{(1+\nu_j)}} \left(M_t^j\right)^{-\frac{1}{(1+\nu_j)}} \left(\frac{C_t^j}{P_t}\right)^{1-\frac{\rho}{1+\nu_j}} \left(\frac{C_t^Q}{P_t} Q_t\right)^{\frac{\rho}{1+\nu_j}} (Q_t)^{\frac{1-\rho}{1+\nu_j}},$$

using the condition $\frac{C_t^Q}{P_t} Q_t = (1-\varphi)Y_t$, and the aggregate production function,

$$\frac{r_t^M}{P_t} = \Gamma \left(1 + \frac{1-\eta}{\nu_j}\right)^{\frac{-\nu_j}{(1+\nu_j)(1-\eta)}} \left(A_t^{jM}\right)^{\frac{\nu_j}{(1+\nu_j)}} \left(M_t^j\right)^{-\frac{1}{(1+\nu_j)}} \left(\frac{C_t^j}{P_t}\right)^{1-\frac{\rho}{1+\nu_j}} ((1-\varphi)Y_t)^{\frac{\rho}{1+\nu_j}} \left(\frac{Y_t^{1/(1-\varphi)}}{K_t^{\varphi/(1-\varphi)}}\right)^{\frac{1-\rho}{1+\nu_j}}$$

$$\frac{r_t^M}{P_t} = \Gamma \left(1 + \frac{1-\eta}{\nu_j}\right)^{\frac{-\nu_j}{(1+\nu_j)(1-\eta)}} (1-\varphi)^{\frac{\rho}{1+\nu_j}} \left(A_t^{jM}\right)^{\frac{\nu_j}{(1+\nu_j)}} \left(M_t^j\right)^{-\frac{1}{(1+\nu_j)}} K_t^{\frac{\varphi(\rho-1)}{(1-\varphi)(1+\nu_j)}} \left(\frac{C_t^j}{P_t}\right)^{\frac{1+\nu_j-\rho}{1+\nu_j}} (Y_t)^{\frac{1-\rho\varphi}{(1+\nu_j)(1-\varphi)}},$$

and similarly for wages,

$$\frac{w_t^j}{P_t} = \Gamma \left(1 + \frac{1-\eta}{\nu_j}\right)^{\frac{-\nu_j}{(1+\nu_j)(1-\eta)}} (1-\varphi)^{\frac{\rho}{1+\nu_j}} \left(A_t^{jH}\right)^{\frac{\nu_j}{(1+\nu_j)}} \left(H_t^j\right)^{-\frac{1}{(1+\nu_j)}} K_t^{\frac{\varphi(\rho-1)}{(1-\varphi)(1+\nu_j)}} \left(\frac{C_t^j}{P_t}\right)^{\frac{1+\nu_j-\rho}{1+\nu_j}} (Y_t)^{\frac{1-\rho\varphi}{(1+\nu_j)(1-\varphi)}}.$$

The hat algebra for these expressions is straightforward,

$$\frac{\hat{r}_{t+1}^m}{\hat{P}_{t+1}} = \left(\hat{A}_{t+1}^{jM}\right)^{\frac{\nu_j}{(1+\nu_j)}} \left(\hat{M}_{t+1}^j\right)^{-\frac{1}{(1+\nu_j)}} K_{t+1}^{\frac{\varphi(\rho-1)}{(1-\varphi)(1+\nu_j)}} \left(\frac{\hat{C}_{t+1}^j}{\hat{P}_{t+1}}\right)^{\frac{1+\nu_j-\rho}{1+\nu_j}} \left(\hat{Y}_{t+1}\right)^{\frac{1-\rho\varphi}{(1+\nu_j)(1-\varphi)}}$$

$$\frac{\hat{w}^j_{t+1}}{\hat{P}_{t+1}} = \left(\hat{A}^{jH}_{t+1}\right)^{\frac{\nu_j}{(1+\nu_j)}} \left(\hat{H}^j_{t+1}\right)^{-\frac{1}{(1+\nu_j)}} K_{t+1}^{\frac{\varphi(\rho-1)}{(1-\varphi)(1+\nu_j)}} \left(\frac{\hat{C}^j_{t+1}}{\hat{P}_{t+1}}\right)^{\frac{1+\nu_j-\rho}{1+\nu_j}} \left(\hat{Y}_{t+1}\right)^{\frac{1-\rho\varphi}{(1+\nu_j)(1-\varphi)}}.$$

We have the unit costs inside these expressions, which itself depends on rental rates and wages. The hat algebra for unit costs is,

$$\frac{\hat{C}^j_{t+1}}{\hat{P}_{t+1}} = \left[(r^M_{t+1}/P_{t+1})^{-\nu_j}(A^{jM}_{t+1})^{\nu_j}\Gamma\left(1+\frac{1-\eta}{\nu_j}\right)^{\frac{\nu_j}{\eta-1}}\left(\frac{C^j_t}{P_t}\right)^{\nu_j} + (w^j_t/P_{t+1})^{-\nu_j}(A^{jH}_t)^{\nu_j}\Gamma\left(1+\frac{1-\eta}{\nu_j}\right)^{\frac{\nu_j}{\eta-1}}\left(\frac{C^j_t}{P_t}\right)^{\nu_j}\right]^{-1/\nu_j}.$$

and using (73),

$$\frac{\hat{C}^j_{t+1}}{\hat{P}_{t+1}} = \left[\left(\frac{\hat{r}^M_{t+1}}{\hat{P}_{t+1}}\right)^{-\nu_j}(\hat{A}^{jM}_{t+1})^{\nu_j}\ \pi^{jM}_t + \left(\frac{\hat{w}^j_{t+1}}{\hat{P}_{t+1}}\right)^{-\nu_j}(\hat{A}^{jH}_{t+1})^{\nu_j}\ \pi^{jH}_t\right]^{-1/\nu_j}$$

We can manipulate the Euler equations to get,

$$\begin{aligned}
\frac{\hat{r}^K_{t+1}}{\hat{P}_{t+1}} &= \frac{\frac{R_{t+1}}{P_{t+1}} - \left(1-\delta^K\right)}{\frac{R_t}{P_t} - (1-\delta^K)} \\
\frac{\hat{r}^M_{t+1}}{\hat{P}_{t+1}} &= \left(\frac{1}{\hat{\xi}^M_{t+1}}\right)\frac{\frac{R_{t+1}}{P_{t+1}} - \left(1-\delta^M\right)\frac{1}{\hat{\xi}^M_{t+2}}}{\frac{R_t}{P_t} - (1-\delta^M)\frac{1}{\hat{\xi}^M_{t+1}}}
\end{aligned}$$

Then, we can express the set of equilibrium conditions under perfect foresight as,

$$\hat{v}^{j}_{t+1} = \Phi^{j}_{t} \left(\frac{\hat{w}^{j}_{t+1}}{\hat{P}_{t+1}}\right)^{1-\gamma} + \left(1-\Phi^{j}_{t}\right)\left[\sum_{\ell=1}^{J} \mu^{j\ell}_{t} \left(\hat{v}^{\ell}_{t+2}\right)^{\frac{\alpha}{1-\gamma}}\right]^{\frac{1-\gamma}{\alpha}} \tag{74}$$

$$\Phi^{j}_{t+1} = \Phi^{j}_{t} \frac{\left[(\hat{w}^{j}_{t+1}\right]^{1-\gamma}}{\hat{v}^{j}_{t+1}} \tag{75}$$

$$\mu^{j\ell}_{t} = \frac{\mu^{j\ell}_{t-1} \left(\hat{v}^{\ell}_{t+1}\right)^{\frac{\alpha}{1-\gamma}}}{\sum_{k=1}^{J} \mu^{j\ell}_{t-1} \left(\hat{v}_{t+1}\right)^{\frac{\alpha}{1-\gamma}}} \tag{76}$$

$$H^{\ell}_{t+1} = (1-\delta)\sum_{j=1}^{J} \mathcal{M}^{j\ell}_{t} H^{j}_{t} + \delta H^{0,\ell}_{t+1} \tag{77}$$

$$\mathcal{M}^{j\ell}_{t} = \left(\frac{\mu^{j\ell}_{t}}{\mu^{j\ell}_{t-1}}\right)^{1-\frac{1}{\alpha}} \mathcal{M}^{j\ell}_{t-1} \tag{78}$$

$$\theta^{\ell}_{t+1} = (1-\delta)\sum_{j=1}^{J} \mu^{j\ell}_{t}\, \theta^{j}_{t} + \delta\theta^{0,\ell}_{t+1} \tag{79}$$

$$\theta^{0,\ell}_{t+1} = \frac{\theta^{0,\ell}_{t} \left(\hat{v}^{\ell}_{t+1}\right)^{\frac{\alpha}{1-\gamma}}}{\sum_{k=1}^{J} \theta^{0,k}_{t} \left(\hat{v}^{k}_{t+1}\right)^{\frac{\alpha}{1-\gamma}}} \tag{80}$$

$$H^{0,\ell}_{t+1} = H^{0,\ell}_{t} \left(\frac{\theta^{0,\ell}_{t+1}}{\theta^{0,\ell}_{t}}\right)^{1-\frac{1}{\alpha}} \tag{81}$$

$$\frac{\hat{w}^{j}_{t+1}}{\hat{P}_{t+1}} = \left(\hat{A}^{jH}_{t+1}\right)^{\frac{\nu_j}{(1+\nu_j)}} \left(\hat{H}^{j}_{t+1}\right)^{-\frac{1}{(1+\nu_j)}} \hat{K}_{t+1}^{\frac{\varphi(\rho-1)}{(1-\varphi)(1+\nu_j)}} \left(\frac{\hat{C}^{j}_{t+1}}{\hat{P}_{t+1}}\right)^{\frac{1+\nu_j-\rho}{1+\nu_j}} \left(\hat{Y}_{t+1}\right)^{\frac{1-\rho\varphi}{(1+\nu_j)(1-\varphi)}} \tag{82}$$

$$\frac{\hat{r}^{m}_{t+1}}{\hat{P}_{t+1}} = \left(\hat{A}^{jM}_{t+1}\right)^{\frac{\nu_j}{(1+\nu_j)}} \left(\hat{M}^{j}_{t+1}\right)^{-\frac{1}{(1+\nu_j)}} \hat{K}_{t+1}^{\frac{\varphi(\rho-1)}{(1-\varphi)(1+\nu_j)}} \left(\frac{\hat{C}^{j}_{t+1}}{\hat{P}_{t+1}}\right)^{\frac{1+\nu_j-\rho}{1+\nu_j}} \left(\hat{Y}_{t+1}\right)^{\frac{1-\rho\varphi}{(1+\nu_j)(1-\varphi)}} \tag{83}$$

$$\frac{\hat{r}^{K}_{t+1}}{\hat{P}_{t+1}} = \frac{\hat{Y}_{t+1}}{\hat{K}_{t+1}} \tag{84}$$

$$\hat{Y}_{t+1} = \left(\hat{K}_{t+1}\right)^{\varphi} \left[\sum_{j=1}^{J} \varrho^{j}_{t} \left[\left(\hat{A}^{jM}_{t+1}\hat{M}^{j}_{t+1}\right)^{\frac{\nu_j}{1+\nu_j}} \pi^{jM}_{t} + \left(\hat{A}^{jH}_{t+1}\hat{H}^{j}_{t+1}\right)^{\frac{\nu_j}{1+\nu_j}} \pi^{jH}_{t}\right]^{\frac{(1+\nu_j)(\rho-1)}{\nu_j\,\rho}}\right]^{\frac{(1-\varphi)\rho}{\rho-1}} \tag{85}$$

$$\varrho^{j}_{t+1} = \frac{\varrho^{j}_{t} \left[\left(\hat{A}^{jM}_{t+1}\hat{M}^{j}_{t+1}\right)^{\frac{\nu_j}{1+\nu_j}} \pi^{jM}_{t} + \left(\hat{A}^{jH}_{t+1}\hat{H}^{j}_{t+1}\right)^{\frac{\nu_j}{1+\nu_j}} \pi^{jH}_{t}\right]^{\frac{(1+\nu_j)(\rho-1)}{\nu_j\,\rho}}}{\sum_{j=1}^{J} \varrho^{j}_{t} \left[\left(\hat{A}^{jM}_{t+1}\hat{M}^{j}_{t+1}\right)^{\frac{\nu_j}{1+\nu_j}} \pi^{jM}_{t} + \left(\hat{A}^{jH}_{t+1}\hat{H}^{j}_{t+1}\right)^{\frac{\nu_j}{1+\nu_j}} \pi^{jH}_{t}\right]^{\frac{(1+\nu_j)(\rho-1)}{\nu_j\,\rho}}} \tag{86}$$

$$\frac{\hat{C}^{j}_{t+1}}{\hat{P}_{t+1}} = \left[\left(\frac{\hat{r}^{M}_{t+1}}{\hat{P}_{t+1}}\right)^{-\nu_j} (\hat{A}^{jM}_{t+1})^{\nu_j}\, \pi^{jM}_{t} + \left(\frac{\hat{w}^{j}_{t+1}}{\hat{P}_{t+1}}\right)^{-\nu_j} (\hat{A}^{jH}_{t+1})^{\nu_j}\, \pi^{jH}_{t}\right]^{-1/\nu_j} \tag{87}$$

$$\pi^{jH}_{t+1} = \frac{\pi^{jH}_{t}(\hat{w}^{j}_{t+1}/\hat{P}_{t+1})^{-\nu_j}\,(\hat{A}^{jH}_{t+1})^{\nu_j}}{\pi^{jH}_{t}(\hat{w}^{j}_{t+1}/\hat{P}_{t+1})^{-\nu_j}\,(\hat{A}^{j}_{t+1})^{\nu_j} + \pi^{m}_{t}(\hat{r}^{m}_{t+1}/\hat{P}_{t+1})^{-\nu_j}\,(\hat{A}^{M}_{t+1})^{\nu_j}} \tag{88}$$

$$\pi^{jM}_{t+1} = \frac{\pi^{jM}_{t}(\hat{r}^{m}_{t+1}/\hat{P}_{t+1})^{-\nu_j}\,(\hat{A}^{jM}_{t+1})^{\nu_j}}{\pi^{jH}_{t}(\hat{w}^{j}_{t+1}/\hat{P}_{t+1})^{-\nu_j}\,(\hat{A}^{jH}_{t+1})^{\nu_j} + \pi^{jM}_{t}(\hat{r}^{m}_{t+1}/\hat{P}_{t+1})^{-\nu_j}\,(\hat{A}^{jM}_{t+1})^{\nu_j}} \tag{89}$$

$$\frac{\hat{r}^K_{t+1}}{\hat{P}_{t+1}} = \frac{\frac{R_{t+1}}{P_{t+1}} - \left(1-\delta^K\right)}{\frac{R_t}{P_t} - \left(1-\delta^K\right)} \tag{90}$$

$$\frac{\hat{r}^M_{t+1}}{\hat{P}_{t+1}} = \left(\frac{1}{\hat{\xi}^M_{t+1}}\right) \frac{\frac{R_{t+1}}{P_{t+1}} - \left(1-\delta^M\right)\frac{1}{\hat{\xi}^M_{t+2}}}{\frac{R_t}{P_t} - \left(1-\delta^M\right)\frac{1}{\hat{\xi}^M_{t+1}}} \tag{91}$$

with $\Phi^j_0 = \frac{(w^j_0/P_0)^{1-\gamma}}{(1-\gamma)v^j_0}$.

It is clear from the previous of equilibrium conditions of the model, expressed in changes, that a very large set of parameters are no longer present, as described in Proposition 4.

If the economy is initialy in a BGP, which is our assumption in the quantitative application, one can pin-down the initial value of the endogenous variable Φ^j_0 given some normalizations.[42] In particular,

$$\Phi^j_0 = 1 - \beta(1-\delta)\Gamma\left(1 - \frac{(1-\gamma)}{\alpha}\right)\mu_{j,j}^{-(1-\gamma)/\alpha}\left(\frac{\mathcal{M}_{jj}}{\Gamma\left(1-1/\alpha\right)\mu_{j,j}^{(1-1/\alpha)}}\right)^{1-\gamma}.$$

B.1 Counterfactuals

We now describe how to conduct counterfactual experiments using the dynamic hat algebra. For this, we assume that at time $t = 1$, a fully unanticipated shock hits the economy in terms of a change in the sequence of TFP or the price of equipment (i.e. $\hat{A}$ or $\hat{\xi}$). The expressions that need a careful adjustment are the dynamic conditions of the workers. We denote with a tilde variables in the counterfactual equilibrium. Note that at time $t = 0$, variables

[42] It is important to highlight that Proposition 4 is more general, and the assumption of an initial BGP is not required. However, outside of the BGP, some additional data is required given that normalizations are no longer valid.

in both cases are the same, as the shock is only know at $t=1$. We then have, for $t=1$,

$$\frac{\tilde{v}_1^j}{v_0^j} = \Phi_0^j \left(\frac{\tilde{w}_1^j/\tilde{P}_1}{w_0^j/P_0}\right)^{1-\gamma} + \left(1-\Phi_0^j\right)\left[\sum_{\ell=1}^{J} \frac{\mu_0^{j\ell}}{\phi_1^\ell}\left(\frac{-\tilde{v}_2^\ell}{-\tilde{v}_1^\ell}\right)^{\frac{\alpha}{1-\gamma}}\right]^{\frac{1-\gamma}{\alpha}} \tag{92}$$

$$\tilde{\mu}_1^{j\ell} = \frac{\frac{\mu_0^{j\ell}}{\phi_1^\ell}\left(\frac{-\tilde{v}_2^\ell}{-\tilde{v}_1^\ell}\right)^{\frac{\alpha}{1-\gamma}}}{\sum_{k=1}^{J}\frac{\mu_0^{j\ell}}{\phi_1^k}\left(\frac{-\tilde{v}_2^\ell}{-\tilde{v}_1^\ell}\right)^{\frac{\alpha}{1-\gamma}}} \tag{93}$$

$$\phi_1^j = \left(\frac{v_1^j/v_0^j}{\tilde{v}_1^j/v_0^j}\right)^{\frac{\alpha}{1-\gamma}} \tag{94}$$

$$\tilde{H}_1^\ell = (1-\delta)\sum_{j=1}^{J}\mathcal{M}_0^{j\ell}\, H_0^j + \delta\tilde{H}_1^{0,\ell} \tag{95}$$

$$\tilde{\mathcal{M}}_1^{j\ell} = \left(\frac{\tilde{\mu}_1^{j\ell}}{\mu_0^{j\ell}}\right)^{1-\frac{1}{\alpha}} \mathcal{M}_0^{j\ell} \tag{96}$$

$$\tilde{\theta}_1^\ell = (1-\delta)\sum_{j=1}^{J}\mu_0^{j\ell}\,\theta_0^j + \delta\tilde{\theta}_1^{0,\ell} \tag{97}$$

$$\tilde{\theta}_1^{0,\ell} = \frac{\frac{\theta_0^{0,\ell}}{\phi_1^\ell}\left(\frac{-\tilde{v}_2^\ell}{-\tilde{v}_1^\ell}\right)^{\frac{\alpha}{1-\gamma}}}{\sum_{k=1}^{J}\frac{\theta_0^{0,k}}{\phi_1^k}\left(\frac{-\tilde{v}_2^k}{-\tilde{v}_1^k}\right)^{\frac{\alpha}{1-\gamma}}} \tag{98}$$

$$\tilde{H}_1^{0,\ell} = H_0^{0,\ell}\left(\frac{\tilde{\theta}_1^{0,\ell}}{\theta_0^{0,\ell}}\right)^{1-\frac{1}{\alpha}} \tag{99}$$

$$\tilde{\Phi}_1^j = \Phi_0^j\,\frac{\left[\frac{\tilde{w}_1^j/w_0^j}{\tilde{P}_1/P_0}\right]^{1-\gamma}}{\tilde{v}_1^j/v_0^j} \tag{100}$$

$$\frac{\tilde{w}_1^j/\tilde{P}_1}{w_0^j/P_0} = \left(\frac{\tilde{A}_1^{jH}}{A_0^{jH}}\right)^{\frac{\nu_j}{(1+\nu_j)}} \left(\frac{\tilde{H}_1^j}{H_0^j}\right)^{-\frac{1}{(1+\nu_j)}} \left(\frac{\tilde{K}_1}{K_0}\right)^{\frac{\varphi(\rho-1)}{(1-\varphi)(1+\nu_j)}} \left(\frac{\tilde{C}_1^j/\tilde{P}_1}{C_0^j/P_0}\right)^{\frac{1+\nu_j-\rho}{1+\nu_j}} \left(\frac{\tilde{Y}_1}{Y_0}\right)^{\frac{1-\rho\varphi}{(1+\nu_j)(1-\varphi)}} \tag{101}$$

$$\frac{\tilde{r}_1^m/\tilde{P}_1}{r_0^m/P_0} = \left(\frac{\tilde{A}_1^{jM}}{A_0^{jM}}\right)^{\frac{\nu_j}{(1+\nu_j)}} \left(\frac{\tilde{M}_1^j}{M_0^j}\right)^{-\frac{1}{(1+\nu_j)}} \left(\frac{\tilde{K}_1}{K_0}\right)^{\frac{\varphi(\rho-1)}{(1-\varphi)(1+\nu_j)}} \left(\frac{\tilde{C}_1^j/\tilde{P}_1}{C_0^j/P_0}\right)^{\frac{1+\nu_j-\rho}{1+\nu_j}} \left(\frac{\tilde{Y}_1}{Y_0}\right)^{\frac{1-\rho\varphi}{(1+\nu_j)(1-\varphi)}} \tag{102}$$

$$\frac{\tilde{r}_1^K/r_0^K}{\tilde{P}_1/P_0} = \frac{\tilde{Y}_1/Y_0}{K_1/K_0} \tag{103}$$

$$\frac{\tilde{Y}_1}{Y_0} = \left(\frac{\tilde{K}_1}{K_0}\right)^{\varphi} \left[\sum_{j=1}^{J} \varrho_0^j \left[\left(\frac{\tilde{A}_1^{jM}}{A_0^{jM}}\frac{\tilde{M}_1^j}{M_0^j}\right)^{\frac{\nu_j}{1+\nu_j}} \pi_0^{jM} + \left(\frac{\tilde{A}_1^{jH}}{A_0^{jH}}\frac{\tilde{H}_1^j}{H_0^j}\right)^{\frac{\nu_j}{1+\nu_j}} \pi_0^{jH}\right]^{\frac{(1+\nu_j)(\rho-1)}{\nu_j\,\rho}}\right]^{\frac{(1-\varphi)\rho}{\rho-1}} \tag{104}$$

$$\tilde{\varrho}_1^j = \frac{\varrho_0^j \left[\left(\frac{\tilde{A}_1^{jM}}{A_0^{jM}}\frac{\tilde{M}_1^j}{M_0^j}\right)^{\frac{\nu_j}{1+\nu_j}} \pi_0^{jM} + \left(\frac{\tilde{A}_1^{jH}}{A_0^{jH}}\frac{\tilde{H}_1^j}{H_0^j}\right)^{\frac{\nu_j}{1+\nu_j}} \pi_0^{jH}\right]^{\frac{(1+\nu_j)(\rho-1)}{\nu_j\,\rho}}}{\sum_{j=1}^{J} \varrho_0^j \left[\left(\frac{\tilde{A}_1^{jM}}{A_0^{jM}}\frac{\tilde{M}_1^j}{M_0^j}\right)^{\frac{\nu_j}{1+\nu_j}} \pi_0^{jM} + \left(\frac{\tilde{A}_1^{jH}}{A_0^{jH}}\frac{\tilde{H}_1^j}{H_0^j}\right)^{\frac{\nu_j}{1+\nu_j}} \pi_0^{jH}\right]^{\frac{(1+\nu_j)(\rho-1)}{\nu_j\,\rho}}} \tag{105}$$

$$\frac{\tilde{C}_1^j/\tilde{P}_1}{C_0^j/P_0} = \left[\left(\frac{\tilde{r}_1^M/\tilde{P}_1}{r_0^M/P_0}\right)^{-\nu_j} \left(\frac{\tilde{A}_1^{jM}}{A_0^{jM}}\right)^{\nu_j} \pi_0^{jM} + \left(\frac{\tilde{w}_1^j/\tilde{P}_1}{w_0^j/P_0}\right)^{-\nu_j} \left(\frac{\tilde{A}_1^{jH}}{A_0^{jH}}\right)^{\nu_j} \pi_0^{jH}\right]^{-1/\nu_j} \tag{106}$$

$$\tilde{\pi}_1^{jH} = \frac{\pi_0^{jH}\left(\frac{\tilde{w}_1^j/\tilde{P}_1}{w_0^j/P_0}\right)^{-\nu_j} \left(\frac{\tilde{A}_1^{jH}}{A_0^{jH}}\right)^{\nu_j}}{\pi_0^{jH}\left(\frac{\tilde{w}_1^j/\tilde{P}_1}{w_0^j/P_0}\right)^{-\nu_j} \left(\frac{\tilde{A}_1^{jH}}{A_0^{jH}}\right)^{\nu_j} + \pi_0^{jM}\left(\frac{\tilde{r}_1^m/\tilde{P}_1}{r_0^m/P_0}\right)^{-\nu_j} \left(\frac{\tilde{A}_1^{jM}}{A_0^{jM}}\right)^{\nu_j}} \tag{107}$$

$$\tilde{\pi}_1^{jM} = \frac{\pi_0^{jM}\left(\frac{\tilde{r}_1^m/\tilde{P}_1}{r_0^m/P_0}\right)^{-\nu_j} \left(\frac{\tilde{A}_1^{jM}}{A_0^{jM}}\right)^{\nu_j}}{\pi_0^{jH}\left(\frac{\tilde{w}_1^j/\tilde{P}_1}{w_0^j/P_0}\right)^{-\nu_j} \left(\frac{\tilde{A}_1^{jH}}{A_0^{jH}}\right)^{\nu_j} + \pi_0^{jM}\left(\frac{\tilde{r}_1^m/\tilde{P}_1}{r_0^m/P_0}\right)^{-\nu_j} \left(\frac{\tilde{A}_1^{jM}}{A_0^{jM}}\right)^{\nu_j}} \tag{108}$$

$$\frac{\tilde{r}_1^K/\tilde{P}_1}{r_0^K/P_0} = \frac{\frac{\tilde{R}_1}{\tilde{P}_1} - \left(1-\delta^K\right)}{\frac{R_0}{P_0} - \left(1-\delta^K\right)} \tag{109}$$

$$\frac{\tilde{r}_1^M/\tilde{P}_1}{r_0^M/P_0} = \left(\frac{1}{\tilde{\xi}_1^M/\xi_0^M}\right)\frac{\frac{\tilde{R}_1}{\tilde{P}_1} - \left(1-\delta^M\right)\frac{1}{\tilde{\xi}_2^M/\tilde{\xi}_1^M}}{\frac{R_0}{P_0} - \left(1-\delta^M\right)\frac{1}{\tilde{\xi}_1^M/\xi_0^M}} \tag{110}$$

where, $\Phi_0^j = \frac{(w_0^j/P_0)^{1-\gamma}}{(1-\gamma)v_0^j}$. Note that, the new variable ϕ_1^j captures the jump in lifetime utility that occurs at time $t=1$, relative to the baseline, due to the unexpected shock.

For all other periods ($t \geq 1$) conditions are very similar to the baseline case:

$$\frac{\tilde{v}^j_{t+1}}{\tilde{v}^j_t} = \tilde{\Phi}^j_t \left(\frac{\tilde{w}^j_{t+1}/\tilde{P}_{t+1}}{\tilde{w}^j_t/\tilde{P}_t}\right)^{1-\gamma} + \left(1-\tilde{\Phi}^j_t\right)\left[\sum_{\ell=1}^{J} \tilde{\mu}^{j\ell}_t \left(\frac{-\tilde{v}^{\ell}_{t+2}}{-\tilde{v}^{\ell}_{t+1}}\right)^{\frac{\alpha}{1-\gamma}}\right]^{\frac{1-\gamma}{\alpha}} \tag{111}$$

$$\tilde{\mu}^{j\ell}_{t+1} = \frac{\tilde{\mu}^{j\ell}_t \left(\frac{-\tilde{v}^{\ell}_{t+2}}{-\tilde{v}^{\ell}_{t+1}}\right)^{\frac{\alpha}{1-\gamma}}}{\sum_{k=1}^{J} \tilde{\mu}^{j\ell}_t \left(\frac{-\tilde{v}^{\ell}_{t+2}}{-\tilde{v}^{\ell}_{t+1}}\right)^{\frac{\alpha}{1-\gamma}}} \tag{112}$$

$$\tilde{H}^{\ell}_{t+1} = (1-\delta)\sum_{j=1}^{J} \tilde{\mathcal{M}}^{j\ell}_t \, \tilde{H}^j_t + \delta \tilde{H}^{0,\ell}_{t+1} \tag{113}$$

$$\tilde{\mathcal{M}}^{j\ell}_{t+1} = \left(\frac{\tilde{\mu}^{j\ell}_{t+1}}{\tilde{\mu}^{j\ell}_t}\right)^{1-\frac{1}{\alpha}} \tilde{\mathcal{M}}^{j\ell}_t \tag{114}$$

$$\tilde{\theta}^{\ell}_{t+1} = (1-\delta)\sum_{j=1}^{J} \tilde{\mu}^{j\ell}_t \, \tilde{\theta}^j_t + \delta \tilde{\theta}^{0,\ell}_{t+1} \tag{115}$$

$$\tilde{\theta}^{0,\ell}_{t+1} = \frac{\tilde{\theta}^{0,\ell}_t \left(\frac{-\tilde{v}^{\ell}_{t+2}}{-\tilde{v}^{\ell}_{t+1}}\right)^{\frac{\alpha}{1-\gamma}}}{\sum_{k=1}^{J} \tilde{\theta}_t^{\,0,k} \left(\frac{-\tilde{v}^{k}_{t+2}}{-\tilde{v}^{k}_{t+1}}\right)^{\frac{\alpha}{1-\gamma}}} \tag{116}$$

$$\tilde{H}^{0,\ell}_{t+1} = \tilde{H}^{0,\ell}_t \left(\frac{\tilde{\theta}^{0,\ell}_{t+1}}{\tilde{\theta}^{0,\ell}_t}\right)^{1-\frac{1}{\alpha}} \tag{117}$$

$$\tilde{\Phi}^j_{t+1} = \Phi^j_t \, \frac{\left[\frac{\tilde{w}^j_{t+1}/\tilde{w}^j_t}{\tilde{P}_{t+1}/\tilde{P}_t}\right]^{1-\gamma}}{\tilde{v}^j_{t+1}/\tilde{v}^j_t} \tag{118}$$

$$\frac{\tilde{w}^j_{t+1}/\tilde{P}_{t+1}}{\tilde{w}^j_t/\tilde{P}_t} = \left(\frac{\tilde{A}^{jH}_{t+1}}{\tilde{A}^{jH}_t}\right)^{\frac{\nu_j}{(1+\nu_j)}} \left(\frac{\tilde{H}^j_{t+1}}{\tilde{H}^j_t}\right)^{-\frac{1}{(1+\nu_j)}} \left(\frac{\tilde{K}_{t+1}}{\tilde{K}_t}\right)^{\frac{\varphi(\rho-1)}{(1-\varphi)(1+\nu_j)}} \left(\frac{\tilde{C}^j_{t+1}/\tilde{P}_{t+1}}{\tilde{C}^j_t/\tilde{P}_t}\right)^{\frac{1+\nu_j-\rho}{1+\nu_j}} \left(\frac{\tilde{Y}_{t+1}}{\tilde{Y}_t}\right)^{\frac{1-\rho\varphi}{(1+\nu_j)(1-\varphi)}} \tag{119}$$

$$\frac{\tilde{r}^m_{t+1}/\tilde{P}_{t+1}}{\tilde{r}^m_t/\tilde{P}_t} = \left(\frac{\tilde{A}^{jM}_{t+1}}{\tilde{A}^{jM}_t}\right)^{\frac{\nu_j}{(1+\nu_j)}} \left(\frac{\tilde{M}^j_{t+1}}{\tilde{M}^j_t}\right)^{-\frac{1}{(1+\nu_j)}} \left(\frac{\tilde{K}_{t+1}}{\tilde{K}_t}\right)^{\frac{\varphi(\rho-1)}{(1-\varphi)(1+\nu_j)}} \left(\frac{\tilde{C}^j_{t+1}/\tilde{P}_{t+1}}{\tilde{C}^j_t/\tilde{P}_t}\right)^{\frac{1+\nu_j-\rho}{1+\nu_j}} \left(\frac{\tilde{Y}_{t+1}}{\tilde{Y}_t}\right)^{\frac{1-\rho\varphi}{(1+\nu_j)(1-\varphi)}} \tag{120}$$

$$\frac{\tilde{r}^K_{t+1}/\tilde{r}^K_t}{\tilde{P}_{t+1}/\tilde{P}_t} = \frac{\tilde{Y}_{t+1}/\tilde{Y}_t}{\tilde{K}_{t+1}/\tilde{K}_t} \tag{121}$$

$$\frac{\tilde{Y}_{t+1}}{\tilde{Y}_t} = \left(\frac{\tilde{K}_{t+1}}{\tilde{K}_t}\right)^{\varphi} \left[\sum_{j=1}^{J} \tilde{\varrho}^j_t \left[\left(\frac{\tilde{A}^{jM}_{t+1}}{\tilde{A}^{jM}_t}\frac{\tilde{M}^j_{t+1}}{\tilde{M}^j_t}\right)^{\frac{\nu_j}{1+\nu_j}} \tilde{\pi}^{jM}_t + \left(\frac{\tilde{A}^{jH}_{t+1}}{\tilde{A}^{jH}_t}\frac{\tilde{H}^j_{t+1}}{\tilde{H}^j_t}\right)^{\frac{\nu_j}{1+\nu_j}} \tilde{\pi}^{jH}_t\right]^{\frac{(1+\nu_j)(\rho-1)}{\nu_j\,\rho}}\right]^{\frac{(1-\varphi)\rho}{\rho-1}} \tag{122}$$

$$\tilde{\varrho}^j_{t+1} = \frac{\tilde{\varrho}^j_t \left[\left(\frac{\tilde{A}^{jM}_{t+1}}{\tilde{A}^{jM}_t}\frac{\tilde{M}^j_{t+1}}{\tilde{M}^j_t}\right)^{\frac{\nu_j}{1+\nu_j}} \tilde{\pi}^{jM}_t + \left(\frac{\tilde{A}^{jH}_{t+1}}{\tilde{A}^{jH}_t}\frac{\tilde{H}^j_{t+1}}{\tilde{H}^j_t}\right)^{\frac{\nu_j}{1+\nu_j}} \tilde{\pi}^{jH}_t\right]^{\frac{(1+\nu_j)(\rho-1)}{\nu_j\,\rho}}}{\sum_{j=1}^{J} \tilde{\varrho}^j_t \left[\left(\frac{\tilde{A}^{jM}_{t+1}}{\tilde{A}^{jM}_t}\frac{\tilde{M}^j_{t+1}}{\tilde{M}^j_t}\right)^{\frac{\nu_j}{1+\nu_j}} \tilde{\pi}^{jM}_t + \left(\frac{\tilde{A}^{jH}_{t+1}}{\tilde{A}^{jH}_t}\frac{\tilde{H}^j_{t+1}}{\tilde{H}^j_t}\right)^{\frac{\nu_j}{1+\nu_j}} \tilde{\pi}^{jH}_t\right]^{\frac{(1+\nu_j)(\rho-1)}{\nu_j\,\rho}}} \tag{123}$$

$$\frac{\tilde{C}^j_{t+1}/\tilde{P}_{t+1}}{\tilde{C}^j_t/\tilde{P}_t} = \left[\left(\frac{\tilde{r}^M_{t+1}/\tilde{P}_{t+1}}{\tilde{r}^M_t/\tilde{P}_t}\right)^{-\nu_j}\left(\frac{\tilde{A}^{jM}_{t+1}}{\tilde{A}^{jM}_t}\right)^{\nu_j}\tilde{\pi}^{jM}_t + \left(\frac{\tilde{w}^j_{t+1}/\tilde{P}_{t+1}}{\tilde{w}^j_t/\tilde{P}_t}\right)^{-\nu_j}\left(\frac{\tilde{A}^{jH}_{t+1}}{\tilde{A}^{jH}_t}\right)^{\nu_j}\tilde{\pi}^{jH}_t\right]^{-1/\nu_j} \tag{124}$$

$$\tilde{\pi}^{jH}_{t+1} = \frac{\tilde{\pi}^{jH}_t\left(\frac{\tilde{w}^j_{t+1}/\tilde{P}_{t+1}}{\tilde{w}^j_t/\tilde{P}_t}\right)^{-\nu_j}\left(\frac{\tilde{A}^{jH}_{t+1}}{\tilde{A}^{jH}_t}\right)^{\nu_j}}{\tilde{\pi}^{jH}_t\left(\frac{\tilde{w}^j_{t+1}/\tilde{P}_{t+1}}{\tilde{w}^j_t/\tilde{P}_t}\right)^{-\nu_j}\left(\frac{\tilde{A}^{jH}_{t+1}}{\tilde{A}^{jH}_t}\right)^{\nu_j} + \tilde{\pi}^{jM}_t\left(\frac{\tilde{r}^m_{t+1}/\tilde{P}_{t+1}}{\tilde{r}^m_t/\tilde{P}_t}\right)^{-\nu_j}\left(\frac{\tilde{A}^{jM}_{t+1}}{\tilde{A}^{jM}_t}\right)^{\nu_j}} \tag{125}$$

$$\tilde{\pi}^{jM}_{t+1} = \frac{\tilde{\pi}^{jM}_t\left(\frac{\tilde{r}^m_{t+1}/\tilde{P}_{t+1}}{\tilde{r}^m_t/\tilde{P}_t}\right)^{-\nu_j}\left(\frac{\tilde{A}^{jM}_{t+1}}{\tilde{A}^{jM}_t}\right)^{\nu_j}}{\tilde{\pi}^{jH}_t\left(\frac{\tilde{w}^j_{t+1}/\tilde{P}_{t+1}}{\tilde{w}^j_t/\tilde{P}_t}\right)^{-\nu_j}\left(\frac{\tilde{A}^{jH}_{t+1}}{\tilde{A}^{jH}_t}\right)^{\nu_j} + \tilde{\pi}^{jM}_t\left(\frac{\tilde{r}^m_{t+1}/\tilde{P}_{t+1}}{\tilde{r}^m_t/\tilde{P}_t}\right)^{-\nu_j}\left(\frac{\tilde{A}^{jM}_{t+1}}{\tilde{A}^{jM}_t}\right)^{\nu_j}} \tag{126}$$

$$\frac{\tilde{r}^K_{t+1}/\tilde{P}_{t+1}}{\tilde{r}^K_t/\tilde{P}_t} = \frac{\frac{\tilde{R}_{t+1}}{\tilde{P}_{t+1}} - \left(1-\delta^K\right)}{\frac{\tilde{R}_t}{\tilde{P}_t} - \left(1-\delta^K\right)} \tag{127}$$

$$\frac{\tilde{r}^M_{t+1}/\tilde{P}_{t+1}}{\tilde{r}^M_t/\tilde{P}_t} = \left(\frac{1}{\tilde{\xi}^M_{t+1}/\tilde{\xi}^M_t}\right)\frac{\frac{\tilde{R}_{t+1}}{\tilde{P}_{t+1}} - \left(1-\delta^M\right)\frac{1}{\tilde{\xi}^M_{t+2}/\tilde{\xi}^M_{t+1}}}{\frac{\tilde{R}_t}{\tilde{P}_t} - \left(1-\delta^M\right)\frac{1}{\tilde{\xi}^M_{t+1}/\tilde{\xi}^M_t}} \tag{128}$$

C Welfare

To characterize welfare in this economy we use the following equilibrium conditions,

$$v^j_t = \frac{(w^j_t/P_t)^{1-\gamma}}{1-\gamma} - \beta\,\Gamma\left(1-\frac{1-\gamma}{\alpha}\right)\left[\sum_{\ell=1}^{J}(-v^\ell_{t+1})^{\frac{\alpha}{1-\gamma}}\ (\tau_{j\ell}\lambda_\ell)^\alpha\right]^{\frac{1-\gamma}{\alpha}} \tag{129}$$

$$\mu^{j\ell}_t = \frac{\left[\lambda_\ell\tau_{j\ell}\left(-v^\ell_{t+1}\right)^{\frac{1}{1-\gamma}}\right]^\alpha}{\sum_{k=1}^J\left[\lambda_k\tau_{jk}\left(-v^k_{t+1}\right)^{\frac{1}{1-\gamma}}\right]^\alpha} \tag{130}$$

Combining these equations we get,

$$v^j_t = \frac{(w^j_t/P_t)^{1-\gamma}}{1-\gamma} + \beta\,\Gamma\left(1-\frac{1-\gamma}{\alpha}\right)\left(\mu^{jj}_t\right)^{-\frac{1-\gamma}{\alpha}}\ \lambda_j^{1-\gamma}\ \tau_{jj}^{1-\gamma}\ v^j_{t+1}. \tag{131}$$

Call $G^v_t = \Gamma\left(1-\frac{1-\gamma}{\alpha}\right)\left(\mu^{jj}_t\right)^{-\frac{1-\gamma}{\alpha}}\ \lambda_j^{1-\gamma}\ \tau_{jj}^{1-\gamma}$. We can further expand this as,

$$v^j_t = \sum_{s=0}^{\infty}\beta^s\left(\prod_{k=0}^{s}G^v_{t+k-1}\right)\frac{(w^j_{t+s}/P_{t+s})^{1-\gamma}}{1-\gamma} \tag{132}$$

Under a counterfactual economy, we have

$$\tilde{v}^j_t = \sum_{s=0}^{\infty}\beta^s\left(\prod_{k=0}^{s}\tilde{G}^v_{t+k-1}\right)\frac{(\tilde{w}^j_{t+s}/\tilde{P}_{t+s})^{1-\gamma}}{1-\gamma} \tag{133}$$

Define δ_j the consumption equivalent measure for the counterfactual economy, such that

$$\tilde{v}^j_t = \sum_{s=0}^{\infty}\beta^s\left(\prod_{k=0}^{s}G^v_{t+k-1}\right)\frac{(\delta_j w^j_{t+s}/P_{t+s})^{1-\gamma}}{1-\gamma} \tag{134}$$

Then, $\delta_j = \left(\tilde{v}_t^j / v_t^j\right)^{1/(1-\gamma)}$.

D Extensions

In this appendix, we show how to extend our baseline framework along various dimensions that are both quantitatively and conceptually relevant.

D.1 Heterogeneous Workers

In addition to their human capital levels, h, and their current occupations, j, workers are heterogenous along two other dimensions. First, workers belong to different groups, indexed by $e = 1, .., E$, defined according to observable characteristics such as the worker's schooling attainment, gender, and possibly race. These worker characteristics are fixed over time. In addition, workers differ in their age, which also governs the horizon for their labor market participation. For tractability, we consider workers that fall into two possible age groups: young and old. These age groups are denoted $a \in \{y, o\}$. We assume a simple form of stochastic aging, i.e. time invariant transitions across age groups: Every period, each a young worker becomes an old worker with probability $0 < \pi \leq 1$, and with probability $1 - \pi$ stays a young worker. Finally, young workers die with probability $0 < \delta^y \leq 1$ every period and old workers die with probability $0 < \delta^o \leq 1$.

In each period, new young workers enter the labor markets. The mass of "newborn" workers is described by a positive measure, $\theta_t^{e,nb,j}$, that defines how they are distributed across fixed characteristics e, and initial occupations j. It is convenient to assume that newborns may be hit by the aging shock upon entry, thus some of them will start the period as old.[43] For newborn workers, the distribution of human capital within their group is $\phi_t^{e,nb,j}(h)$. Similarly, at any point in time, the mass of workers of different characteristics in the economy and the distribution of human capital within each of these groups are $\theta_t(e, a, j)$ and $\phi_t^{e,a,j}(h)$, respectively.

All workers take as given the vector of unitary wages $w = [w(1), w(2), ...w(J)]$ for all the J occupations. The individual characteristics e and the age a of the worker determine the pecuniary and non-pecuniary costs, $\tau_{j,\ell}^{e,a}$ and $\chi_{j,\ell}^{e,a}$, of switching occupations, as well as the parameters that govern the probabilities for labor market opportunities $\lambda_\ell^{e,a}$ for all occupations. Aside from these forms of heterogeneity, the structure of the dynamic choices of workers is exactly the same as in the basic model, as we explain in detail now.

The value function for young, and old workers depends on their type e, and are respectively given by:

$$
\begin{aligned}
V^{e,y}(j,h,\epsilon) &= \frac{[w(j)h]^{1-\gamma}}{1-\gamma} + \beta^y \max_\ell \left\{ \chi_{j,\ell}^{e,y} \pi E_{\epsilon'|y,e} \left[V^{e,o}\left[\ell, h\tau_{j,\ell}^{e,y}\epsilon_\ell,\, \epsilon'\right]\right] + \chi_{j,\ell}^{e,y}(1-\pi) E_{\epsilon'|y,e}\left[V^{e,y}\left[\ell, h\tau_{j,\ell}^{e,y}\epsilon(\ell), \epsilon'\right]\right]\right\}, \\
V^{e,o}(j,h,\epsilon) &= \frac{[w(j)h]^{1-\gamma}}{1-\gamma} + \beta^o \max_\ell \left\{ \chi_{j,\ell}^{e,o} E_{\epsilon'|o,e}\left[V^{e,o}\left[\ell, h\tau_{j,\ell}^{e,o}\epsilon_\ell, \epsilon'\right]\right]\right\},
\end{aligned}
$$

where $E_{\epsilon'|a,e}$ denotes the expectation with respect to labor market opportunities ϵ', conditional on the age and type of workers, and $\beta^y = \beta(1-\delta^y)$ and $\beta^o = \beta(1-\delta^o)$ are the effective discount factors which take into account the death probability.

Following the same steps as in the basic model, we can define by $v^{e,y}(j)$, $v^{e,o}(j)$ as the expectation of the normalized ($h = 1$) value of a worker in group e attached to occupation j. These values are given by the Bellman equations

$$
\begin{aligned}
v^{e,y}(j) &= \frac{[w(j)]^{1-\gamma}}{1-\gamma} + \beta^y E_{\epsilon|y,e} \max_\ell \left\{ [\pi v^{e,o}(\ell) + (1-\pi) v^{e,y}(\ell)]\; \chi_{j,\ell}^{e,y} \, [\tau_{j,\ell}^{e,y}\epsilon_\ell]^{1-\gamma} \right\}, \\
v^{e,o}(j) &= \frac{[w(j)]^{1-\gamma}}{1-\gamma} + \beta^o E_{\epsilon|o,e} \max_\ell \left\{ v^{e,o}(\ell)\, \chi_{j,\ell}^{e,o} \, [\tau_{j,\ell}^{e,y}\epsilon_\ell]^{1-\gamma} \right\},
\end{aligned}
$$

Our quantitative model imposes the assumption that the vector of market opportunities ϵ for all workers is

[43] As we discuss next, this assumption is convenient since all transition matrices μ and $\mathcal{M}$ will have all elements strictly positive and the version of the Perron-Frobenius theorem for positive matrices holds. Nonetheless, we can relax this assumption, in which case these matrices will have some elements equal to zero. We would need to check that these matrices are irreducible and apply the version of the theorem for this case.

distributed according to a multidimensional extreme value distribution, which we allow to vary across the type e and age a of the worker. Specifically, we assume that each ϵ_ℓ is independently distributed Frechet with age- and type-dependent scale parameters $\lambda_\ell^{e,a} > 0$ but common curvature $\alpha > 1$. The expected labor market growth factors for a worker of group e and age a are

$$\Phi^{e,a}(j) \equiv \Gamma\left(1 - \frac{1-\gamma}{\alpha}\right)\left[\sum_{\ell=1}^{J}(\chi_{j,\ell}^{e,a})^{\frac{\alpha}{1-\gamma}}\left(\tau_{j\ell}^{e,a}\lambda_\ell^{e,a}\right)^{\alpha}\right],$$

and define $\bar{\Phi}^{e,a} \equiv \max_j \{\Phi^{e,a}(j)\}$ the highest of those growth factors for each of the age and type groups. Finally, redefine the bounds on flow utilities $\underline{u}$ and $\bar{u}$ to be the uniform bounds across all a and e. Then, under the conditions that $\beta^a\bar{\Phi}^{e,a} < 1$ for all e, a, we can find finite lower and upper bounds $\underline{v}$, $\bar{v}$ that apply to all the normalized expected values $v^{e,a}(j)$.

The following corollary extends the characterization of the individual worker's problem of Theorem 3 to be applied in our quantitative model.

Corollary 1 ***Individual Problems, Extended Model.*** *Assume that the shocks $\epsilon = [\epsilon_1, ...\epsilon_\ell, ...\epsilon_J]$ are distributed Frechet with shape parameter $\alpha > 1$ and scale parameters $\lambda_\ell^{e,a} > 0$ for all ages a and types e. Assume also that w is strictly positive and that $\beta^a\bar{\Phi}^{e,a} < 1$ for all e, a. Then:* ***(i)*** *If $0 \le \gamma < 1$, the expected values $v(j)$ for $j = 1, ..., J$ solve the fixed point problem*

$$\begin{aligned} v^{e,y}(j) &= \frac{(w_j)^{1-\gamma}}{1-\gamma} + \beta^y\Gamma\left(1-\frac{1-\gamma}{\alpha}\right)\left[\sum_{\ell=1}^{J}\left[\pi v^{e,o}(\ell) + (1-\pi)v^{e,y}(\ell)\right]^{\frac{\alpha}{1-\gamma}}(\chi_{j,\ell}^{e,y})^{\frac{\alpha}{1-\gamma}}\left(\tau_{j\ell}^{e,y}\lambda_\ell^{e,y}\right)^{\alpha}\right]^{\frac{1-\gamma}{\alpha}}. \\ v^{e,o}(j) &= \frac{(w_j)^{1-\gamma}}{1-\gamma} + \beta^o\Gamma\left(1-\frac{1-\gamma}{\alpha}\right)\left[\sum_{\ell=1}^{J}\left[v^{e,o}(\ell)\right]^{\frac{\alpha}{1-\gamma}}(\chi_{j,\ell}^{e,o})^{\frac{\alpha}{1-\gamma}}\left(\tau_{j\ell}^{e,o}\lambda_\ell^{e,o}\right)^{\alpha}\right]^{\frac{1-\gamma}{\alpha}} \end{aligned}$$

A solution $v^{e,a} \in [0, \bar{v}]^J$ for these BE exists and is unique. Moreover, the proportion of workers switching from occupation j to occupation ℓ is given by:

$$\begin{aligned} \mu_{j\ell}^{e,y} &= \frac{\left[\pi v^{e,o}(\ell) + (1-\pi)v^{e,y}(\ell)\right]^{\frac{\alpha}{1-\gamma}}(\chi_{j,\ell}^{e,y})^{\frac{\alpha}{1-\gamma}}\left(\tau_{j\ell}^{e,y}\lambda_\ell^{e,y}\right)^{\alpha}}{\sum_{k=1}^{J}\left[\pi v^{e,o}(k) + (1-\pi)v^{e,y}(k)\right]^{\frac{\alpha}{1-\gamma}}(\chi_{j,k}^{e,y})^{\frac{\alpha}{1-\gamma}}\left(\tau_{jk}^{e,y}\lambda_k^{e,y}\right)^{\alpha}} \\ \mu_{j\ell}^{e,o} &= \frac{\left[v^{e,o}(\ell)\right]^{\frac{\alpha}{1-\gamma}}(\chi_{j,\ell}^{e,o})^{\frac{\alpha}{1-\gamma}}\left(\tau_{j\ell}^{e,o}\lambda_\ell^{e,o}\right)^{\alpha}}{\sum_{k=1}^{J}\left[v^{e,o}(k)\right]^{\frac{\alpha}{1-\gamma}}(\chi_{j,k}^{e,o})^{\frac{\alpha}{1-\gamma}}\left(\tau_{jk}^{e,o}\lambda_k^{e,o}\right)^{\alpha}}. \end{aligned}$$

(ii) *If $\gamma > 1$ the expected values $v(j)$ for $j = 1, ..., J$, solve the fixed point problem*

$$\begin{aligned} v^{e,y}(j) &= \frac{(w_j)^{1-\gamma}}{1-\gamma} - \beta^y\Gamma\left(1-\frac{1-\gamma}{\alpha}\right)\left[\sum_{\ell=1}^{J}\left[\pi(-v^{e,o}(\ell)) + (1-\pi)(-v^{e,y}(\ell))\right]^{\frac{\alpha}{1-\gamma}}(\chi_{j,\ell}^{e,y})^{\frac{\alpha}{1-\gamma}}\left(\tau_{j\ell}^{e,y}\lambda_\ell^{e,y}\right)^{\alpha}\right]^{\frac{1-\gamma}{\alpha}}. \\ v^{e,o}(j) &= \frac{(w_j)^{1-\gamma}}{1-\gamma} - \beta^o\Gamma\left(1-\frac{1-\gamma}{\alpha}\right)\left[\sum_{\ell=1}^{J}\left[(-v^{e,o}(\ell))\right]^{\frac{\alpha}{1-\gamma}}(\chi_{j,\ell}^{e,o})^{\frac{\alpha}{1-\gamma}}\left(\tau_{j\ell}^{e,o}\lambda_\ell^{e,o}\right)^{\alpha}\right]^{\frac{1-\gamma}{\alpha}} \end{aligned}$$

A solution $v \in [\underline{v}, 0]^J$ for this BE exists and is unique. Moreover, the proportion of workers that switch from *occupation j to occupation ℓ is given by:*

$$\begin{aligned}
\mu_{j\ell}^{e,y} &= \frac{\left[\pi(-v^{e,o}(\ell)) + (1-\pi)(-v^{e,y}(\ell))\right]^{\frac{\alpha}{1-\gamma}} \ (\chi_{j,\ell}^{e,y})^{\frac{\alpha}{1-\gamma}} \ \left(\tau_{j\ell}^{e,y}\lambda_{\ell}^{e,y}\right)^{\alpha}}{\sum_{k=1}^{J}\left[\pi(-v^{e,o}(k)) + (1-\pi)(-v^{e,y}(k))\right]^{\frac{\alpha}{1-\gamma}} \ (\chi_{j,k}^{e,y})^{\frac{\alpha}{1-\gamma}} \ \left(\tau_{jk}^{e,y}\lambda_{k}^{e,y}\right)^{\alpha}} \\
\mu_{j\ell}^{e,o} &= \frac{\left[(-v^{e,o}(\ell))\right]^{\frac{\alpha}{1-\gamma}} \ (\chi_{j,\ell}^{e,o})^{\frac{\alpha}{1-\gamma}} \ \left(\tau_{j\ell}^{e,o}\lambda_{\ell}^{e,o}\right)^{\alpha}}{\sum_{k=1}^{J}\left[(-v^{e,o}(k))\right]^{\frac{\alpha}{1-\gamma}} \ (\chi_{j,k}^{e,o})^{\frac{\alpha}{1-\gamma}} \ \left(\tau_{jk}^{e,o}\lambda_{k}^{e,o}\right)^{\alpha}}.
\end{aligned}$$