The Long and Lasting Effects of Government R&D

Jason Hall*

PRELIMINARY AND INCOMPLETE: PLEASE DO NOT SHARE OR CITE

January 27, 2026

Abstract

Government R&D spending has two primary effects. Firstly, it directly leads to the creation of new ideas. Secondly, it encourages individuals to invest in costly, specific, and irreversible scientific human capital. I study the latter by building a tractable model of multi-sector economic growth that features heterogeneous knowledge workers and heterogeneous firms. Despite the substantial heterogeneity on both sides, the model cleanly aggregates. I then develop an extended version of the model featuring correlation in scientific skills across individuals. I calibrate the model to restricted use data from the United States Census Bureau and National Center for Science and Engineering Statistics on the earnings and mobility patterns of individuals receiving doctorates from United States institutions from 2000 to 2017. I then use the calibrated model to understand the impact of government subsidies for different scientific fields on aggregate growth and welfare.

*University of Minnesota and Federal Reserve Bank of Minneapolis. hall2160@umn.edu

I thank my advisors Anmol Bhandari, Kjetil Storesletten, Hannes Malmberg, Todd Schoellman, and Kyle Herkenhoff for detailed comments on this and all other projects. I have also benefited from extensive conversations with Marco Bassetto and Simon Mongey. Many thanks to Loukas Karabarbounis and all students in the Growth and Public and Macro-Labor workshops at the University of Minnesota for their comments and suggestions as well. I have no conflicts of interest to disclose.

1 Introduction

Growth in the United States and much of the developed world is driven by the accumulation of new knowledge. This knowledge can be generated by the activities of many actors: universities, firms, or government research labs to name the most aware. Regardless of where production takes place, the central ingredient of this activity is specialized human capital. This specialization comes in the form of two specific sets of choices by individuals. The first is the choice of being a researcher or not. Virtually all modern growth models feature this choice. The second choice pertains to the specific type of knowledge that individuals choose to acquire. Virtually no modern growth models feature this choice.

In this paper, I fill this gap empirically and theoretically. Empirically, I leverage historical variation in federal support for different universities and fields to estimate, the elasticity of PhD supply to federal support for R&D. Motivated by these empirical results in aggregate data, I then leverage restricted-use micro data from the Survey of Earned Doctorates that provides a near-census of PhD recipients in the United States from the year 2000 to 2017, matched into the Longitudinal Employer Household Dynamics dataset of the United States Census Bureau and patent data. This provides me earnings, employment histories, and measure of innovative output for approximately **X%** of PhD Recipients in the United States from 2000 to 2017.[1]

This dataset permits me to document a number of new facts about the contributions and earnings profiles of PhD recipients in the United States. Firstly, I document that PhD recipients constitute a vastly disproportionate share of aggregate innovation, with PhD recipients in my sample constituting roughly **X%** of patenting by individuals in the United States in the period running from 2000 to 2022. This average prospensity, however, masks substantial heterogeneity by field, with some fields having much wider differences while other major fields having much lower. Consistent with this, PhD recipients are substantially more likely to be involved with R&D activities than the general population or even college educated individuals. Secondly, I use a revealed preference argument to construct measures of the knowledge content of different industries. Specifically,

[1] Anything in red is pending disclosure review. Please see Section 7 for time-table to completion.

I leverage a linkage between the SED and LEHD to construct measures of exposures of different industries to different PhD fields and find substantial heterogeneity in the knowledge content of different industries. Thirdly, I estimate, cohort-by-cohort, earnings profiles by major field. The earnings profiles I estimate are informative as to the profile of human capital accumulation by individuals in different field, which is relevant for measuring the impact of federal research support on aggregate growth and welfare. The earnings profiles themselves, however, are contaminated by age, time, and cohort effects. I leverage the panel dimension of my data to overcome this issue, crucially allowing me to identify the shape of human capital accumulation on a major field basis.

Armed with these empirical inputs, I develop a novel endogenous growth model. My model substantially enriches the baseline model of researcher labor supply, and features costly and irreversible investment into specific human capital, life-cycle accumulation of experience by researchers and heterogeneous knowledge inputs by sectors. I use the earnings profiles estimated in the prior step as an input into the estimation of a system of cross-price elasticities of PhD labor supply, which capture the extent to which changes in the path of wages of e.g. biologists change the labor supply of e.g. mathematicians. These short-run elasticities, coupled with the experience profiles I estimate, characterize the long-run labor supply effects of short-run fluctuations in federal funding.

To estimate the effects of proposed cuts to federal R&D spending, I further enrich the model to include direct federal support for research, and use existing estimates of the magnitude of spillovers from federal research to discipline the magnitude of these spillovers. This allows me to engage in a set of counterfactual exercises. I then examine the importance of the NIH Doubling for the increase in doctors with degrees in biological sciences and health sciences. I find that holding fixed NIH funding levels at 1998 levels would have lead to $K\%$ fewer graduate students and lowered productivity growth in sectors making use of biology by nearly $k\%$. My second counterfactual examines the extent to which one-time large increases generate smaller quantity responses than smaller but prolonged increases. Specifically, I consider an exercise in which I take the 10 billion dollars allocated to the NIH as part of the American Recovery and Reinvestment Act, and then

spread that 10 billion dollar one-time increase over 10 years. Next, I study the effect of the proposed cuts to federal funding for R&D. I find that - in a comparison of steady-states - the subsidies would change long-run growth by nearly $X\%$ at a welfare cost of $Y\%$. Including the transition changes this welfare cost to $Z\%$.

In the final section of the paper, I engage in an optimal policy exercise where inclusive of the transition, I solve a Ramsey problem.

1.1 Related Literature

This paper builds on the voluminous literature that studies innovation and endogenous growth. In my model, as in the canonical endogenous growth models of Romer (1990) and Aghion and Howitt (1992), ideas are the engine of aggregate growth, and ideas are generated by researchers. A number of papers have emphasized the importance of the allocation of talent for economic growth (see e.g. Murphy et al., 1991; Aghion et al., 2017; Bell et al., 2019; Hsieh et al., 2019). Most related to this paper is recent work in Akcigit et al. (2025); the authors develop an endogenous growth model that features an endogenous choice of research versus production occupation, R&D capital and endogenous training periods to study the degree of complementarity between R&D and education policies. They then calibrate the model to a set of novel facts about characteristics of PhDs derived from Danish administrative data. Relative to that work, my data allow me to study the choices of what to study by those individuals who choose to get a doctoral degree, and to study the relative importance of those choices. Accordingly, my model further endogenized the choice of PhD field, and features demand for PhD fields that varies across sectors. Both features allow me to quantify the effectiveness of subsidies to PhD support by field and the efficacy of these across different sectors. A small literature, most notably Jones (2009), has also endogenized this choice but my model is to the best of my knowledge the first quantitative model to do so.

A parallel but related strand of literature in the economics of science has studied characteristics of graduate choice and described the outcomes of graduates after the PhD (Buffington et al., 2016;Chang et al., 2019;Zolas et al., 2015;Jiang et al., 2023). Existing papers have mostly relied

on linkages between UMETRICS data and either the Longitudinal Household Earnings Dynamics data (LEHD) or tax forms, with the exception of work in Foster et al. (2023a) and Foster et al. (2023b), who, like me, make use of linkages between the Survey of Earned Doctorates and the Longitudinal Household Earnings Dynamics data, although they restrict their focus to economics PhDs. The usage of the SED provides me dramatically larger sample sizes than existing papers and allows me to see the patenting contributions of the vast majority of new PhD recipients in the United States in the years from 2001 to 2017.

Finally, I relate to the literature on the returns to government R&D. On the micro-side, Azoulay et al. (2019) and Myers and Lanahan (2022) exploit windfall designs to recover the returns to NIH and DOE funding respectively. Both articles find significant spillover effects from federal R&D. Babina et al. (2023) study the effects of federal funding shocks and find that shocks to government funding are partially offset by private funding, but that this tilts innovation towards patentable outputs. Finally, Tham et al. (2024) study the impact of NIH funding delays on graduate education and find that funding interruptions increase the probability of not working in the United States by 3 percentage points. More directly related to my study are existing studies on the macro impact of government R&D, as exemplified in Dyevre (2023) and Fieldhouse and Mertens (2025). Dyevre (2023) exploits two parallel designs - one a historical shift-share, the other a judge design - to estimate an elasticity of productivity growth to government R&D of 0.17. Fieldhouse and Mertens (2025) construct a sequence of narrative shocks to the major federal R&D accounts, and find the returns to government R&D that range from 140 to 210 percent. Relative to this literature, I focus on the impact of government R&D support on the allocation of scientific talent.

Having reviewed the relevant literature, I now summarize the rest of the paper. In Section 2 I first describe the data and motivating facts of my analysis. Section 3 introduces the model. Section 4 describes the estimation procedure for the model and evaluates the fit. Section 5 uses the calibrated model to evaluate several counterfactuals, while Section 6 solves for optimal policy in the model economy before concluding.

2 Motivation

2.1 Government R&D Funding Increases the Number of PhDs

The central premise of this paper is that an important channel through which government funding effects innovation outcomes is the training and creation of new researchers. Accordingly, one would like to know if government funding actually affects the number of PhD students, and if so how much. Simple regressions of outcomes like the number of PhD recipients on government funding would likely conflate general changing in the taste for education as the United States becomes richer from the expansion of fellowship slots or changes in the expected future wages. To purge those concerns in the data, I first leverage changes in NIH funding that began in the early 1990s and culminated in the "NIH doubling period" that ran from fiscal year 1998 to fiscal year 2003. Specifically, I leverage variation in exposure to the rise in NIH funding by comparing the number of graduates of PhD programs in the biomedical sciences with others. Letting c denote a college, f a major field, and t calendar time, I estimate the following via a Poisson Difference-in-Difference design:

$$PhD_{cft} = \alpha_{cf} + \gamma_t + \eta_t \times \mathbf{1}[Biomedical_f] + \epsilon_{cft} \tag{1}$$

Figure 1 plots the results of this exercise. To emphasize, I do not claim that this is a causal estimate of the effect of government funding on relative PhD enrollment. One would naturally be concerned that the 1990s saw a large increase in venture capital interest in biotech, which would independently change demand for biomedical PhDs (Lazonick and Tulum, 2011). Such estimates would lead one to overstate the effect of government R&D on number of new PhDs. Accordingly, I take a second approach using the narrative shocks of Fieldhouse and Mertens (2025), cumulated over time. I match each of the budgetary accounts in the Fieldhouse and Mertens (2025) data to individual major fields.[2] I additionally weight the narrative shocks by the average share of

[2]The budgetary accounts are the R&D accounts for the Departments of Defense, Health and Human Services, Energy (both defense and non-defense), National Aeronautics and Space Administration, and National Science Foundation.

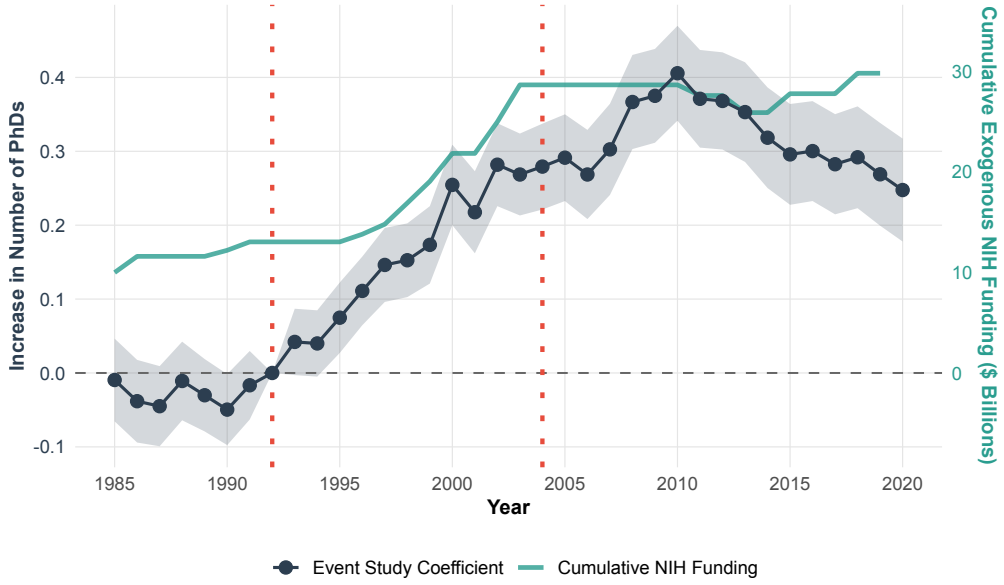


Figure 1: Dynamic Difference-in-Difference plot estimates of η_t. Standard Errors are clustered at major field level. Includes all $R1$ universities and $4 - year$ specialist universities in engineering and medical services. Time series comes from exogeneous narrative funding shocks in Fieldhouse and Mertens (2025).

that account's research funding that accrues to universities, non-profits, or FFRDCs operated by universities in the Survey of Federal Funds for R&D.[3]

Let a denote a budgetary account, ω_a the share of funding that is paid out to to universities, non-profits, and university run FFRDCs, and define the funding by agencies a to a given field by:

$$Fund_{f,t} = \sum_{a \in A_f} \omega_a Fund_{a,t}$$

Letting Δ_s^t denote the time difference operator from s to t, I estimate the following specification for horizons $h \in \{-4, \ldots, -1, 1, \ldots, 12\}$:

$$\ln(PhD_{f,t+h}) - \ln(PhD_{f,t}) = \beta_0 + \beta_1 \left(\frac{\Delta_{t-6}^t Fund_f}{Fund_{f,t-6}} \right) + \epsilon_{ft} \tag{2}$$

[3]Timing limitations for this dataset means that these shares are only available from 1999 to 2019, and I use the average of these shares over the earlier time period. To the extent that these shares vary over time, such variation would show up as measurement error and attenuate the estimates of β_1.

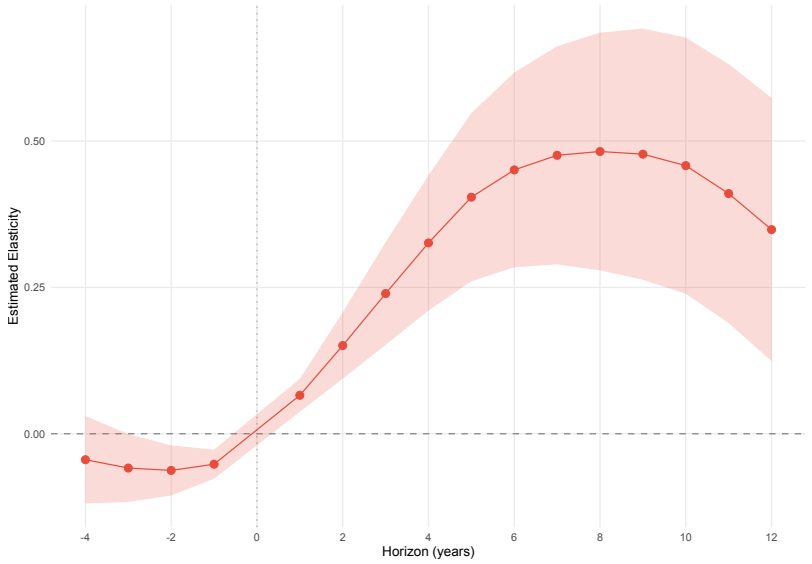


Figure 2: Results of long-difference specification. Sample period starts in 1970. Changes in funding are calculated as log difference between t and $t-6$ to align with length of PhD programs,. Agencies included are the Departments of Defense, Energy, Health and Human Services, NASA, and the NSF. Funding measures are scaled by share of agency's research funding that accrues to universities, university operated FFRDCs or non-profits in the years from 1999 to 2019.

Specification (2) is a dynamic long-difference specification that uses the differences in the cumulative level of funding between $t-6$ and t to capture the dynamic effect of sustained increases in government funding for science on different science fields. I estimate these regressions starting in 1970.[4] My preferred specification implies that a 10% increase in funding for university research of a given field increases the number of doctoral recipients at affiliated fields by roughly 4.5% at peak. This effect peaks from years 7 to 9 before declining in follow-up periods. This is consistent with the duration of PhD programs and lends credence to the plausibility of the estimate of the elasticity of government spending to changes in government R&D.[5]

A natural question is what mechanism drives the positive association between doctoral recipients and government funding for universities. In particular, one might want to distinguish

[4] I choose to neglect the early sample period from 1958 because much of the early budgetary shocks are driven by the space program. NASA funding is generally less targeted at universities than comparable accounts at the NIH and NSF, and the space program in particular featured very large non-research components, as documented in Kantor and Whalley (2025).

[5] Appendix A.1 includes a variety of robustness checks, including varying the timing used for growth rates, extending the sample to include years prior to 1970, adding field fixed effects, estimating the regressions without weighting for university funds, and plotting the raw data at each horizon for the baseline.

between direct support for graduate students via additional funding for fellowships or teaching assistantships and an increase in derived demand from increases in research funds available to faculty. Appendix A.2 shows that the largest mechanism of support for doctoral graduate students is research assistantships, consistent with a view of graduate support as being driven by derived demand.

2.2 PhD Contributions to Innovation

If government funding increases the number of PhDs, a natural next question to ask is what are the contributions of these researchers to aggregate growth? To answer these questions in a descriptive sense, I rely on matched administrative data from the United States Census Bureau and National Center for Science and Engineering Statistics. Firstly, I make use of the Longitiduginal Household Earnings Dynamics (LEHD) from the United States Census Bureau. This dataset is constructed from unemployment insurance filings and covers more than 95% of employment in the United States. My sample of the LEHD includes 28 states.[6] Additional details on the datasets used are available in Appendix G. Secondly, I use the restricted access version of the Survey of Earned Doctorates (SED). The SED is a survey sent to all recipients of PhDs from accredited institutions in the United States. Response rates are typically greater than 90% and the survey is augmented with additional information from university administrators to help mitigate the effect of non-response. As a consequence, this survey should be considered a census of all PhD recipients in the United States in a given year. I match this data in the LEHD to recover the earnings of individuals in these states. Finally, I make use of internal Census copies of the PatentsView database, matched to the LEHD and SED via the crosswalk in Akcigit and Goldschlag (2023).[7]

This matched administrative data allows me to comment on the contribution of doctoral recipients to aggregate innovation. Perhaps unsurprisingly, PhD recipients in my sample account

[6] These are AZ, CA, CO, CT, DE, ID, IN, IA, KS, KY, LA, MA, ME, MD, MN, NE, NV, NJ, NM, ND, OK, OR, PA, SC, SD, UT, VA, WA, and WY.

[7] Note that both the patent data and earnings data are matched independently by PIK, so while my earnings data is limited to only individuals working in the aforementioned states, the patent information covers the contributions from people working in states that I do not have LEHD data for.

Figure 3: Relative probability of a patent in the top 10% of the citation distribution: **PENDING DISCLOSURE REVIEW.**

for a substantial portion of aggregate innovation: nearly X% of all patents and Y% percentage of all inventors in XXXX. Their contributions are concentrated in patent classes X, Y, Z. PhDs also account for a significant portion of patents at the top of citation distribution. Figure 3 plots the probability of a PhD inventor in the top 10% of the citation distribution relative to the population of all inventors.

2.3 Human Capital Profiles of PhD Recipients

A natural next question to ask is what do the life-cycle earnings profiles of individuals with doctorates look like. These profiles are important because they reflect human capital accumulation on the part of researchers, which matters for the timeline in which the effects of government policy are observed. To measure the human capital profiles of individuals with doctorates in a given field, I use the matched LEHD-SED data on earnings and PhD fields. I assume that earnings features an additive structure of the form:

$$\ln(Earn_{iaft}) = \underbrace{\ln(\alpha_{cf})}_{\text{Cohort}} + \underbrace{\ln(\eta_{ft})}_{\text{Skill Price}} + \underbrace{\ln(\theta_{af})}_{\text{Age}} + \epsilon_{iaft} \tag{3}$$

As is well-known, age-time-cohort models suffer from a problem of perfect multicollinearity that renders (3) unestimable without additional restrictions. I follow a strategy similar to Lagakos et al. (2018) in assuming that there is an age at which human capital growth rates are known. In particular, I assume that human capital growth plateaus 20 years post-PhD. This strategy is subject to two caveats. The first is that imposing that human capital is known at more points in time provides additional over-identifying restrictions at the cost of greater risk of misspecification. The second is that I am limited in the duration of available data to me given my sample consists of doctoral recipients beginning in the year 2000, which, restricts my options.

Figure 4: Human Capital Profiles by Major Field: **PENDING DISCLOSURE REVIEW**

Figure 5: Payroll expenditure share by NAICS 2 digit sector and field: **PENDING DISCLOSURE REVIEW**

The estimated plots of human capital are reproduced in Figure 4. The estimation process also generates a sequence of estimated skill prices for each of the major fields and non-research labor. These estimates of (the growth of) skill prices over time will be crucial inputs for the estimation of the structural model.

2.4 Different Industries use Different Fields

The final piece of empirical evidence I assemble is a collection of the payroll-share of PhDs by field (indexed by k) for different NAICS sectors (indexed by s). These expenditure shares reflect the share of total payroll expense to doctoral recipients by individuals with doctorates in a major field. Specifically, this is a matrix $\boldsymbol{\alpha}$ whose (s, k) entry is given by:

$$\boldsymbol{\alpha}_{sk} = \frac{Payroll_k}{\sum_k Payroll_{sk}} \tag{4}$$

.

I represent this matrix via heatmap in Figure 5 below.

3 Model

To understand the relative contribution of different research fields to growth requires heterogeneity in two dimensions. The first dimension is heterogeneity in the knowledge content of sectoral innovation production functions. The second dimension is heterogeneity in the supply of such types of skills. To study this setting, I augment a multi-sector version of a standard quality ladder model with an occupational choice over production and heterogeneous types of research labor.

3.1 Preferences, Endowments and Demographics

Time is continuous and index by t. At any instant, there is a unit mass of individuals that are endowed with a unit of time that they supply inelastically and that have log preferences[8] over flow consumption of the final good, so that:

$$u(c) = \ln(c) \tag{5}$$

Individuals stochastically die and are replaced at a rate δ, so that aggregate population is constant at some mass $\bar{H}$, which I will normalize to 1. I index an individual's age by τ. As individuals die, they are replaced by new individuals who, at birth, draw a $K+1$ dimensional vector of skills from a multivariate Frechet distribution with correlation function given by

$$G(x_0, x_1, \ldots, x_k) = \sum_{f \in F} \left(\sum_{k \in K} (\omega_{fk} x_k)^{\frac{1}{1-\rho_f}} \right)^{1-\rho_f} \tag{6}$$

The marginals of each type of skill are given by:

$$F_k(z) = \exp\left\{-T_k z^{-\theta}\right\} \tag{7}$$

This cross-nested constant elasticity of substitution correlation function, drawn from Lind and Ramondo (2023), is capable of representing substantially more complex patterns of substitution

[8] This is not essential.

than versions commonly used in the literature. To see this, note that, for instance, the correlated CES of Bryan and Morten (2019) corresponds with the case that there is a single nest with correlation ρ. This correlation function generates additional flexibility in measuring the elasticity of substitution in three ways. The first of these is that it allows for the possibility of multiple nests. This corresponds with the Frechet version of the familiar nested logit model. The second margin of flexibility is heterogeneous correlation within each of the nests, so that, for instance, skills in STEM fields may be more tightly correlated than those embedded in non-STEM fields. The third margin is that it allows a single choice to lie in multiple nests. This flexibility resolves the somewhat unattractive problem of independence of irrelevant alternatives. In this context, IIA would require that substitution patterns from other fields to computer science are proportional to the existing share of those fields, which does not comport with one's regular intuition that, for instance, Mathematics and Computer Science are closer substitutes than English and Computer Science.

Once individuals choose their type at birth, they begin to accumulate additional skill in that field according to some learning by doing function $h_k(\tau)$. I will later parameterize this learning-by-doing function, but it is sufficient to assume that $h_k(\tau)$ satisfies two properties:

1. $\forall k \in K \; h_k(\tau) \in \mathcal{C}^2(\mathbb{R}^+)$
2. $\lim_{\tau \to \infty} h_k(\tau)$ exists for all k and is bounded above by a scalar $\bar{h}$.

The resultant aggregate supply of effective units of labor is a consequence of the interplay between the heterogeneous (realized) initial distributions of skill and the demographic structure of the population.

3.2 Technology

Here I describe the production environment of this economy. The final good is produced as a Cobb-Douglas aggregate across a number of sectors.

$$Y_s(t) = \prod_{s \in S} Y_s^{\beta_s} \tag{8}$$

Each of the sectoral goods are produced as a Cobb-Douglas aggregate of a unit-mass of products.

$$Y_s = \exp\left\{ \int_0^1 \ln\big(\sum_{i \in I_j} y_{ist}(j)\big) dj \right\} \tag{9}$$

Within each product, producers make use of a labor technology that is linear in labor:

$$y_{ist}(j) = z_{ist}(j) l_{ist}(j)$$

As in Klette and Kortum (2004), each product line is endowed with a technology that generates new ideas. This generating an arrival rate of new innovations by aggregating across a collection of different types of labor in a Cobb-Douglas manner, where the weights vary according to the sector, to capture the notion that different sectors use different kinds of knowledge workers. When an idea arrives, it generates a productivity improvement that jumps the frontier productivity in a given product line in that sector by a factor λ. The obsolete technology then diffuses to the fringe.

$$a_{st} = \frac{1}{\epsilon_s^{\frac{1}{\gamma}}} \left(\prod_{k \in K} H_{s,k}^{\alpha_{s,k}} \right)^{\frac{1}{\gamma}} \tag{10}$$

There are two aggregate resource feasibility constraints that must hold. Firstly, aggregate output must be equal to aggregate consumption.

$$C_t = Y_t \tag{11}$$

Secondly, the aggregate supply of each type of labor must be equal to it's aggregate use. Letting $z_t(\tau, \kappa)$ denote the distribution of z conditional on being age τ and type k at time t

$$\int_{\Omega_\tau} \int_{\Omega_z} h_k(\tau) z_{kt}(\tau, k) \delta e^{-\delta \tau} dz d\tau = \sum_{s \in S} H_{s,k} \tag{12}$$

3.3 The Decentralized Economy

Having described the physical environment of my model, I now turn to describing the decentralization that I study in this paper. Workers in my economy are named by four states: their asset holdings a, their productivity at their chosen field z, their age τ, and their chosen field k. As in Blanchard (1985) and Yaari (1965), individuals have access to annuity markets. These annuity markets provide an increase in return equal to δ in exchange for forfeiting all assets at death. As a consequence, the instaneous return on an asset position a_t is given by $r_t + \delta$. This implies that at any point a worker faces the following flow budget constraint:

$$\dot{a}_t = (r + \delta)a_t + w_{kt} z h_k(\tau) - c_t \tag{13}$$

Individuals face the natural borrowing limit. Neglecting the mention of this borrowing limit and substituting for the law of motion, the Hamilton-Jacobi-Bellman equation that characterizes a worker's problem conditional on a field k is given by the following expression:

$$(\rho + \delta)V_t(a, z, \tau; k) - \dot{V}_t = \max_c \left\{ \ln(c_t) + \frac{\partial V_t}{\partial a}\left[(r_t + \delta)a_t + w_{kt} z h_k(\tau) - c\right] + \frac{\partial V_t}{\partial \tau}\right\} \tag{14}$$

Here an individual is subject to a natural borrowing limit which I have neglected to write. All individuals in my economy are born with zero assets. As a consequence, the problem of a newborn can be written as:

$$N_t(\mathbf{z}) = \arg\max_{k \in K} \left\{ V_t(0, z_k, 0, k) \right\} \tag{15}$$

There are four groups of producers who solve problems. The first of these is the competitive final goods aggregator. Their problem is given by

$$\max_{Y_s} \prod_{s \in S} Y_s^{\beta_s} - \sum_{s \in S} P_s Y_s \tag{16}$$

Sectoral aggregators take as given the residual demand that they face from the final goods producer. Their problem is also static and given by

$$\max_{y_s(j)} P_s \exp\left(\int_0^1 \ln(y_s(j))dj\right) - \int_0^1 p_s(j)y_s(j)dj \tag{17}$$

Within a given product line, firms take as given the residual demand curve they face, as well as their and their competitors' productivity. They compete 'a-la Bertrand, which, combined with the assumption that varieties are perfect substitutes within a product line implies that the most productive firm is the only one that will operate in equilibrium and further that they will engage in limit pricing. Let $\bar{z}_s(j)$ denote the productivity of the second-most productive firm. After imposing the limit pricing condition, the problem of the most productive firm in a product line is thus given by:

$$\begin{aligned} &\max_{y_s(j)} p_s(j)y_s(j) - w_{0t}\frac{y_s(j)}{z_s(j)} \\ &\text{S. t. } p_s(j) = \frac{w_{0t}}{\bar{z}_s(j)} \end{aligned} \tag{18}$$

Having specified the static payoff of a firm with a given physical production technology, I now specify the dynamic problem of a firm, which is simply a collection of product lines. Firms choose a collection of research labor at any given time, taking as given the wages of each type of research labor, to generate an arrival rate of innovations I_{st}. This problem can be separated into two parts. In the first part, a firm picks the cost minimizing mix of inputs as a function of desired innovation rate. After picking this cost-minimizing mix, they then pick the optimal innovation rate. After solving the cost-minimization problem, which can be found in Appendix **??**, a firm's problem problem is given by the following Hamilton-Jacobi-Bellman equation, where π_{st} denote time t flow profits of a firm in sector s, and $\tilde{w}_{st}$ denotes the usual Cobb-Douglas price index:

$$(r_t + \iota_{st} - g_t)V_{st} - \dot{V}_{st} = \max_{I_s} \pi_{st} - \epsilon_{st}\tilde{w}_{st}I_s^{\gamma} + I_s V_{St} \tag{19}$$

The last thing remaining before defining a competitive equilibrium is to specify the asset market structure that determines the aggregate interest rate. In my economy, households save in the mutual fund that owns the world. The return on this mutual fund is given by r_t. As this is the only instrument in my economy, asset market clearing thus requires that the aggregate value of firms in the economy must be equal to the aggregate savings of all households.

$$\sum_{s \in S} V_{st} = \mathbf{A}_t \tag{20}$$

We are now in a position to define a competitive equilibrium.

A **dynamic competitive equilibrium** is a collection of prices $r_t, \{p_{sjt}\}, \{w_{kt}\}$, allocations $C_t, \{y_{sjt}\}, \{H_{kst}\}, \{\mathcal{L}_{st}\}$, firm values $\{V_s t\}$, mutual fund value V_t^{MF} and growth rates g_t such that:

1. Worker field choices are consistent with (15) and consumption-savings choices are consistent with (14).
2. Intermediate firms solve (18) and firm values are consistent with (19).
3. $\{H_{kst}\}$ solves the cost-minimization problem of firms.
4. Goods markets clear according to (11).
5. Labor markets clear so that (12) holds.
6. Asset markets clear so that (20) holds.

As a corollary of this, I also state the definition of a balanced growth path.

A **balanced growth path** is a dynamic competitive equilibrium satisfying the property that aggregate consumption C_t, aggregate output Y_t, and wages for each type of labor w_{kt} all grow at a constant rate g.

3.4 Characterizing the Competitive Equilibrium

Having defined an equilibrium in this environment, I now characterize the competitive equilibrium. I first prove a series of aggregation results. The first of these results is a variant of the standard static aggregation result that says that this economy admits a static aggregate production function with the usual Cobb-Douglas TFP term, and a production wage that is known in closed form.

Proposition 1. *Let s denote a sector and t denote calendar time. This economy admits a static aggregation of the production of physical output, with values given by:*

$$Y_t = A\mathcal{Z}_t L^P \tag{21}$$

$$\mathcal{Z}_t = \sum_{s\in S} \beta_s \ln(Z_{st}) \tag{22}$$

$$g_t = \ln(\lambda) \sum_{s\in S} \beta_s \iota_{st} \tag{23}$$

$$w_{0t} = \frac{1}{\lambda} A\mathcal{Z}_t \tag{24}$$

Where ι_{st} denotes the endogenous rate of new idea creation in sector s at time t.

Proof. See Appendix B.1. ■

Having characterized the static allocation, I can now characterize the problem of a worker over time. This problem is given in (14).

Lemma 1. *Given the initial condition that newborns start with zero financial wealth, so that $a = 0$, the solution to* (14) *after normalization is given by:*

$$v_t(\hat{a}, z, k, \tau) = B_t + \frac{1}{\rho + \delta} \ln(\hat{a} + \psi_k(t, \tau) z) \tag{25}$$

Where B_t is a time-dependent adjustment term and $\psi_k(t, \tau)$ is the value of a unit of skill at time t

given a worker is of age τ, which can be written in integral form as:

$$\psi_k(t,\tau) = \int_0^\infty \exp\left\{ -\int_0^s r_{t+u} + \delta - g_{t+u} du \right\} \hat{w}_{k,t+s} h_k(\tau + s) ds \tag{26}$$

Furthermore, the consumption policy given $(\hat{a}, z, k, \tau)$ is given by:

$$c_t = (\rho + \delta)\big(\hat{a} + \psi_k(t,\tau) z\big) \tag{27}$$

Proof. See Appendix C.2. ■

This characterization of the normalized problem in turn allows one to write a clean characterization of the problem of the newborn.

Corollary 1. *Let k^* be the solution to the problem given by:*

$$\arg\max_k v_t(0, z_k, k, 0)$$

Then, the share of individuals choosing k is known in closed form as:

$$\mathbb{P}_k = \sum_f \left\{ \frac{\left[\sum_{k'} (\omega_{fk'} T_{k'} \psi_{k'}(t,0)^\theta)^{\frac{1}{1-\rho_f}}\right]^{1-\rho_f}}{\sum_f \left[\sum_{k'} (\omega_{fk'} T_{k'} \psi_{k'}(t,0)^\theta)^{\frac{1}{1-\rho_f}}\right]^{1-\rho_f}} \times \frac{(\omega_{fk} T_k \psi_k(t,0)^\theta)^{\frac{1}{1-\rho_f}}}{\sum_{k'} (\omega_{fk'} T_{k'} \psi_{k'}(t,0)^\theta)^{\frac{1}{1-\rho_f}}} \right\}$$

Further, the effective flow rate of age 0 skill is given by:

$$\mathcal{E}_{kt} = h_k(0) T_k^{\frac{1}{\theta}} \mathbb{P}_k^{1-\frac{1}{\theta}} \Gamma\left(1 - \frac{1}{\theta}\right) \left[\sum_f \left\{ \left[\sum_{k'} (\omega_{fk'} T_{k'} \psi_{k'}(t,0)^\theta)^{\frac{1}{1-\rho_f}} \right]^{1-\rho_f} \frac{(\omega_{fk} T_k \psi_k(t,0)^\theta)^{\frac{1}{1-\rho_f}-1} \omega_{fk}}{\sum_{k'} (\omega_{fk'} T_{k'} \psi_{k'}(t,0)^\theta)^{\frac{1}{1-\rho_f}}} \right\} \right]^{\frac{1}{\theta}}$$

Proof. See Appendix D. ■

These two results allow me to state one of the central results of the paper. This result allows me to characterize the aggregate dynamics of type k human wealth $\mathbf{H}_{kt}$ and effective labor supply $\mathbf{L}_{kt}$ as a system of coupled ordinary differential equations.

Proposition 2. *Suppose that $h_k(\tau) = \kappa_k - (\kappa_k - 1)e^{-\lambda_k \tau}$. Then aggregate labor supply of type k workers at time t can be written as:*

$$\mathbf{L}_{kt} = \kappa_k \tilde{\mathbf{L}}_{kt} - (\kappa_k - 1)\hat{\mathbf{L}}_{kt} \tag{28}$$

Where $\tilde{\mathbf{L}}_{kt}$ and $\hat{\mathbf{L}}_{kt}$ satisfy:

$$\dot{\tilde{\mathbf{L}}}_{kt} = \delta \mathcal{E}_{kt} - \delta \tilde{\mathbf{L}}_{\mathbf{kt}} \tag{29}$$

$$\dot{\hat{\mathbf{L}}}_{kt} = \delta \mathcal{E}_{kt} - (\delta + \lambda_k)\hat{\mathbf{L}}_{\mathbf{kt}} \tag{30}$$

One can write aggregate human wealth of type k workers as:

$$\mathbf{H}_{kt} = \kappa_k \tilde{\mathbf{H}}_{kt} \tilde{\mathbf{L}}_{kt} - (\kappa_k - 1)\hat{\mathbf{H}}_{kt} \hat{\mathbf{L}}_{kt}$$

Where $\tilde{\mathbf{H}}_{kt}$ and $\hat{\mathbf{H}}_{kt}$ satisfy:

$$\dot{\tilde{\mathbf{H}}}_{kt} = (r_t + \delta - g_t)\tilde{\mathbf{H}}_{kt} - \hat{w}_{kt} \tag{31}$$

$$\dot{\hat{\mathbf{H}}}_{kt} = (r_t + \delta - g_t + \lambda_k)\hat{\mathbf{H}}_{kt} - \hat{w}_{kt} \tag{32}$$

Time t value of age τ workers of type k is given by:

$$\psi_k(t, \tau) = \kappa_k \tilde{\mathbf{H}}_{kt} - (\kappa_k - 1)e^{-\lambda_k \tau}\hat{\mathbf{H}}_{kt} \tag{33}$$

Furthermore, aggregate consumption is given by:

$$\mathbf{C}_t = (\rho + \delta)\big(\mathbf{A}_t + \sum_k \mathbf{H}_{kt}\big) \tag{34}$$

Proof. See Appendix C.2 ■

The essence of Proposition 2 is that despite the rich heterogeneity in the earnings profiles and

endogenous selection, aggregate labor supply and human wealth are only a function of 4 state variables for each type. Furthermore, the selection model, depends only on a known function of these state variables across the k types. The dynamics of each of these state variables are governed by simple linear ordinary differential equations. The small dimensionality of this system is a key component in making the model computationally tractable.[9] The final step to complete the characterization of the model is to derive the equations for innovation rates by sector, which are derived from the optimality conditions of the idea producing firms.

Lemma 2. *The optimal innovation rates are ι_{st} are given by the solution to:*

$$\iota_s = \left[\frac{\gamma \epsilon_{st} \tilde{w}_{st}}{V_{st}}\right]^{\frac{1}{1-\gamma}} \tag{35}$$

Equilibrium firm values are given by the following ordinary differential equation:

$$(r_t + \iota_{st})V_{st} - \dot{V}_{st} = \pi_{st} + \left(1 - \frac{1}{\gamma}\right)\iota_{st} V_{st} \tag{36}$$

Furthermore, the labor market clearing condition of research labor of type k can be written as:

$$w_{kt}\mathbf{L}_{kt} = \frac{1}{\gamma}\sum_s \alpha_{s,k} V_{st}\iota_{st} \tag{37}$$

Proof. See Appendix B.2. ■

Having characterized the competitive equilibrium allocation, I now briefly summarize my computational approach. A complete collection of the set of equilibrium conditions and the approach that I take to this solve this differential-algebraic system of equations can be found in Appendix F.[10]

[9] This tractability extends to the addition of exponential training periods, which can be handled with the addition of a single further state variable, or additional exponential terms to change the shape of the human capital accumulation function.

[10] NB: This appendix is not yet completed.

4 Estimation

Estimation of the model pertains in two steps. In the first step, I set some parameters externally. I set ρ, the rate of time preference to 0.03. I set the death rate to 0.025 to correspond with a 40 year career. I set the matrix of Cobb-Douglas weights in the innovation production function to match the estimated values presented in Figure 5. I set the vector of Cobb-Douglas sectoral weights in the final goods aggregator, $\boldsymbol{\beta}$ to match value-added shares in the decade period prior to 2000. I set λ to 0.1 to generate a markup of 10%. I set γ, the curvature in the innovation production function to 2.

The process by which I estimate the parameters of the labor supply system and the productivity of innovative firms. The first of these objects is $\boldsymbol{\Omega}$, the vector of allocation weights to each nest of each choice. The second is the vector of within-nest correlation parameters $\boldsymbol{\rho}$. These two sets of parameters govern the correlation structure of latent skill. The remaining parameters to identify are the vector of Frechet scale weights $\mathbf{T}$, which I have assumed to be constant over time, and the productivity process of the innovative firms, which generates the demand for knowledge labor in the model. As is standard in these models, the latent skill distribution is inherently unitless, and hence is not independently identified of wages. Accordingly, I set $T_k = 1$ for all k.

To understand the sources of variation that identify the correlation structure, it is informative to derive the elasticity of shares with respect to changes in the payoffs. Appendix E.1 shows that the own- and cross-price elasticities are given by, respectively:

$$\frac{\partial \ln(\mathbb{P}_k)}{\partial \ln(\psi_k)} = \theta(1 - \mathbb{P}_k) \tag{38}$$

$$\frac{\partial \ln(\mathbb{P}_k)}{\partial \ln(\psi_i)} = \frac{\theta}{\mathbb{P}_k}\left[\sum_f \left(\frac{\rho_f}{\rho_f - 1}\mathbb{P}_{i|f}\mathbb{P}_{k|f}\mathbb{P}^f\right) - \mathbb{P}_i\mathbb{P}_k\right] \tag{39}$$

Where recall that choice probabilities can be written as:

$$\mathbb{P}_k = \sum_f \left\{ \underbrace{\frac{\left[\sum_{k'} \left(\omega_{fk'} \psi_{k'}^{\theta}\right)^{\frac{1}{1-\rho_f}}\right]^{1-\rho_f}}{\sum_f \left[\sum_{k'} \left(\omega_{fk'} \psi_{k'}^{\theta}\right)^{\frac{1}{1-\rho_f}}\right]^{1-\rho_f}}}_{\equiv \mathbb{P}^f} \times \underbrace{\frac{\left(\omega_{fk} \psi_k^{\theta}\right)^{\frac{1}{1-\rho_f}}}{\sum_{k'} \left(\omega_{fk'} \psi_{k'}^{\theta}\right)^{\frac{1}{1-\rho_f}}}}_{\equiv \mathbb{P}_{k|f}} \right\}$$

To understand the patterns in the data that serve to identify $\mathbf{\Omega}$ and $\boldsymbol{\rho}$, it is helpful to imagine a simple case with four choices, so that $|K| = 4$ and two latent factors, so that $|F| = 2$.[11] Suppose that two of the choices lie different nests. In that case, the initial skill draws are uncorrelated, and the effective cross-price elasticity between the two is zero is simply $\partial \ln(\mathbb{P}_k)/\partial \ln(\psi_i) = -\theta \mathbb{P}_i$. Contrast this with another choice which lies entirely in the same nest as choice i. Calling this choice j delivers a cross-price elasticity $\partial \ln(\mathbb{P}_k)/\partial \ln(\psi_j) = \frac{\theta}{\mathbb{P}_i}\left[\left(\frac{\rho_1}{\rho_1 - 1}\mathbb{P}_{i|1}\mathbb{P}_{j|1}\mathbb{P}^1\right) - \mathbb{P}_i\mathbb{P}_j\right]$. Since $\rho_f \in [0, 1]$, this elasticity is more negative - the same-size wage change will induce weakly larger substitution, holding fixed the ex-ante proportions of i and j. Note that as ρ becomes larger, the strength of this effect becomes more negative, with the limit as $\rho \to 1$ converging to perfect correlation within the nest. Although the correlation structure is more complicated when the allocation weights ω_{kf} are not discrete 1 and 0, the logic carries through. The pattern of substitution carries information about the allocation matrix $\mathbf{\Omega}$, while the strength of substitution carries information about the correlation within each of the nests $\boldsymbol{\rho}$.

There are two challenges in estimating the labor supply system in this model. The first one is mechanical: the estimation of cross-price elasticities requires variation in relative prices. Such variation does not exist along a balanced growth path, which by definition has constant relative prices. Accordingly, I estimate the model along a transition, where I assume that the economy is in the initial balanced growth path in the year 2000 and then a new path for the productivity process of each sector is revealed. I select these paths for a given guess of the correlation parameters to match the estimated skill prices shown in 4.[12]

[11] If one additionally sets $\rho_1 = \rho_2$ then this corresponds with the case of nested Frechet. Additional examples of correlation structures nested by CNCES correlation can be found in Appendix E.2.

[12] These figures are not in 3 but will be put in an appendix because of quantity. I need to have one plot for each PhD major field and production.

Figure 6: Estimated Allocation Matrix Ω for each PhD Field.

Table 1: Calibration Parameters

Parameter		Value	Moment	Model	Data
Externally calibrated					
ρ	Discount rate	0.030	Standard Value		
δ	Death rate	0.025	40-year career		
$\boldsymbol{\alpha}$	Matrix of Cobb-Douglas weights				
$\boldsymbol{\beta}$	Vector of sectoral value-added weights				
Internally calibrated					
λ	Step Size	1.052	Growth rate from Fernald TFP series	%	%
Policy Parameters					
τ_k	Research subsidy for k	0.200			

The second issue plaguing estimation in this setting is that typical variation used to identify Roy models in settings such as mine makes use of switcher designs.[13] This is unfortunately not feasible in my setting: individuals very rarely attain multiple PhDs. To overcome this challenge, I leverage the panel structure of my data. In my model, individuals make choices about their doctoral field given a sequence of forward looking prices. The sequence of prices will change over time as productivity changes. Hence, given the sequence of prices, differences in the choice probabilities of cohorts over time identifies the parameters of the labor supply system. I implement this by optimizing over Ω and $\boldsymbol{\rho}$ to match the observed pattern of shares across cohorts over time. Figure 6 plots the estimated allocation matrix.

To understand how $\boldsymbol{\rho}$ is identified, observe that given an allocation matrix, each element of $\boldsymbol{\rho}$ controls the strength of substitution within a given nest. Accordingly, given the measured change in flow prices, a higher ρ implies a greater quantity response. Of course, this estimation occurs jointly with the estimation of the allocation matrix. Table 1 summarizes the calibrated parameters of the system, apart from those in terms of the labor supply system.

[13] See e.g. Bryan and Morten (2019), Grigsby (2022), Fan (2025).

4.1 Validation

I plan to validate the model by matching the reduced form estimates in Figure 2.

5 Counterfactuals

I consider three counterfactuals:

1. No NIH Doubling
2. Spread the NIH ARRA funding over 10 years but keep PDV the same.
3. Proposed Trump cuts to science funding.

6 Optimal Constrained Policy

In this section, I study optimal policy in an environment where the planner faces a constraint that they can raise at most 10% of GDP in taxes.

6.1 Conclusion

7 Time-Table to Completion

1. January 30 to February 6: submission of disclosure packet to Census for existing empirical inputs to model (see statistics in red) in Sections 1 and 2.

2. February 7 to February 20: solve modified version of model with government R&D demand sector. This is a relatively straightforward modification of existing code infrastructure that solves the model without a government R&D sector.

3. February 21 to March 13: estimation of structural model. I have confirmed in Monte-Carlos with comparable sample size that the estimation strategy works for a case where individuals are myopic.

4. March 14 to March 21: evaluation of counterfactuals and optimal policy. The code for this is written, and thus can be run immediately after estimaton is completed.

5. March 22 onwards: additional extensions including explicitly augmenting the model with training periods. The math for this is worked out already in Appendix C.2.

References

Aghion, Philippe, Ufuk Akcigit, Ari Hyytinen, and Otto Toivanen (2017) "The social origins of inventors,"Technical report, National Bureau of Economic Research.

Aghion, Philippe and Peter Howitt (1992) "A MODEL OF GROWTH THROUGH CREATIVE DESTRUCTION," *Econometrica*, 60 (2), 323–351.

Akcigit, Ufuk and Nathan Goldschlag (2023) "Measuring the characteristics and employment dynamics of US inventors,"Technical report, National Bureau of Economic Research.

Akcigit, Ufuk, Jeremy Pearce, and Marta Prato (2025) "Tapping into talent: Coupling education and innovation policies for economic growth," *Review of Economic Studies*, 92 (2), 696–736.

Azoulay, Pierre, Joshua S Graff Zivin, Danielle Li, and Bhaven N Sampat (2019) "Public R&D investments and private-sector patenting: evidence from NIH funding rules," *The Review of economic studies*, 86 (1), 117–152.

Babina, Tania, Alex Xi He, Sabrina T Howell, Elisabeth Ruth Perlman, and Joseph Staudt (2023) "Cutting the innovation engine: how federal funding shocks affect university patenting, entrepreneurship, and publications," *The Quarterly Journal of Economics*, 138 (2), 895–954.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen (2019) "Who becomes an inventor in America? The importance of exposure to innovation," *The Quarterly Journal of Economics*, 134 (2), 647–713.

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

Bryan, Gharad and Melanie Morten (2019) "The aggregate productivity effects of internal migration: Evidence from Indonesia," *Journal of Political Economy*, 127 (5), 2229–2268.

Buffington, Catherine, Benjamin Cerf, Christina Jones, and Bruce A Weinberg (2016) "STEM training and early career outcomes of female and male graduate students: Evidence from UMETRICS data linked to the 2010 census," *American Economic Review*, 106 (5), 333–338.

Chang, Wan-Ying, Wei Cheng, Julia Lane, and Bruce Weinberg (2019) "Federal funding of doctoral recipients: What can be learned from linked data," *Research Policy*, 48 (6), 1487–1492.

Dyevre, Arnaud (2023) "Public R&D spillovers and productivity growth,"Technical report, Working paper. Eaton, J., & Kortum, S.(1999). International technology ....

Fan, Tianyu (2025) "The Labor Market Incidence of New Technologies," *arXiv preprint arXiv:2504.04047*.

Fieldhouse, Andrew and Karel Mertens (2025) "The returns to government R&D: Evidence from US appropriations shocks."

Foster, Lucia, Erika McEntarfer, and Danielle H Sandler (2023a) "Diversity and Equity in Labor Market Outcomes for Economists," *Journal of Economics, Race, and Policy*, 6 (4), 246–257.

——— (2023b) "Early Career Paths of Economists inside and outside of Academia," *Journal of Economic Perspectives*, 37 (4), 231–250.

Grigsby, John R (2022) *Skill heterogeneity and aggregate labor market dynamics* (w30052): National Bureau of Economic Research.

Hsieh, Chang-Tai, Erik Hurst, Charles I Jones, and Peter J Klenow (2019) "The allocation of talent and us economic growth," *Econometrica*, 87 (5), 1439–1474.

Jiang, Xuan, Joseph Staudt, and Bruce A Weinberg (2023) "A tale of two fields? stem career outcomes,"Technical report, National Bureau of Economic Research.

Jones, Benjamin F (2009) "The burden of knowledge and the "death of the renaissance man": Is innovation getting harder?" *The Review of Economic Studies*, 76 (1), 283–317.

Kantor, Shawn and Alexander Whalley (2025) "Moonshot: Public R&D and growth," *American Economic Review*, 115 (9), 2891–2925.

Klette, Tor Jakob and Samuel Kortum (2004) "Innovating firms and aggregate innovation," *Journal of political economy*, 112 (5), 986–1018.

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman (2018) "Life cycle wage growth across countries," *Journal of Political Economy*, 126 (2), 797–849.

Lazonick, William and Öner Tulum (2011) "US biopharmaceutical finance and the sustainability of the biotech business model," *Research Policy*, 40 (9), 1170–1187.

Lind, Nelson and Natalia Ramondo (2023) "Trade with correlation," *American Economic Review*, 113 (2), 317–353.

Murphy, Kevin M, Andrei Shleifer, and Robert W Vishny (1991) "The allocation of talent: Implications for growth," *The quarterly journal of economics*, 106 (2), 503–530.

Myers, Kyle R and Lauren Lanahan (2022) "Estimating spillovers from publicly funded R&D: Evidence from the US Department of Energy," *American Economic Review*, 112 (7), 2393–2423.

Press, William H (2007) *Numerical recipes 3rd edition: The art of scientific computing*: Cambridge university press.

Romer, Paul M (1990) "Endogenous technological change," *Journal of political Economy*, 98 (5, Part 2), S71–S102.

Tham, Wei Yang, Joseph Staudt, Elisabeth Ruth Perlman, and Stephanie D Cheng (2024) "Scientific talent leaks out of funding gaps," *arXiv preprint arXiv:2402.07235*.

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

Zolas, Nikolas, Nathan Goldschlag, Ron Jarmin et al. (2015) "Wrapping it up in a person: Examining employment and earnings outcomes for Ph. D. recipients," *Science*, 350 (6266), 1367–1371.

A Additional Empirical Results

This appendix contains additional empirical results, including robustness for the long-difference specifications.

A.1 Robustness for Long-Differences

This subsection provides a number of plots that examine the robustness of the findings in the main text about the elasticity of PhD labor supply to changes in government funding. The first of these is robustness to the sample period. Including the early period of the sample generates substantially lower estimates of the estimated elasticity. The composition of the early tilts strongly towards NASA funding for the space program, which was mostly not provided to universities. Because the weighting scheme relies on data from 1999 onwards, changes in the weight scheme will show up as measurement error in the measured growth rate of field support error, introducing attenuation bias into the estimated elasticity.

A.2 Federal Funding for Graduate Students

Figure 13 documents that the primary mechanism of support for most graduate students in most STEM fields is research assistantships. The major exceptions are mathematics, where teaching assistantships dominate, and some less research oriented fields like psychology.

One might also be worried that the marginal source of support for a graduate student is not the same as the average. In this case, increases in federal funding might increase the number of PhD students through a mechanism over than derived demand. Figure 14 suggests that this is not the case. The majority of the increase in the number of supported graduate students during the NIH doubling period came from additional research funds, not from additional fellowship or teaching grants, which saw relatively muted increases.

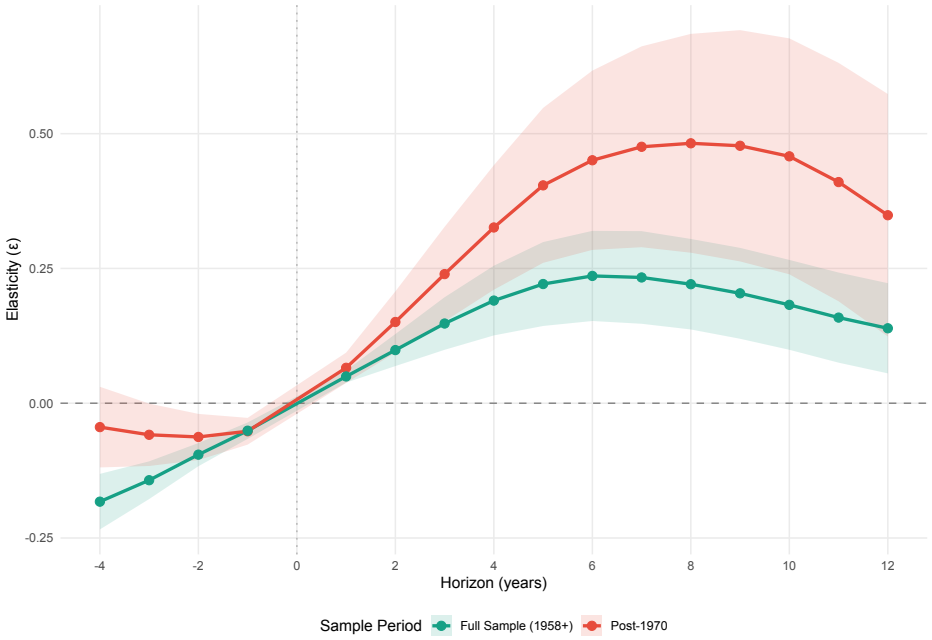


Figure 7: Estimates of Elasticity of Number of PhDs to Change in Government Funding: Robustness for Sample Period. Government funding increases are measured as the difference in level of cumulated FM shocks between $t-k$ and t. Weights are measured using the average share of agency R&D obligations to universities, university operated FFRDCs, and not for profits relative to total yearly R&D obligations from the 1999-2019 in the Survey of Federal Funds for Research and Development. Weights: NIH=71%,NSF = 87%,NASA = 13%,DOE = 11%, DOD = 5%. Standard errors are clustered by field.

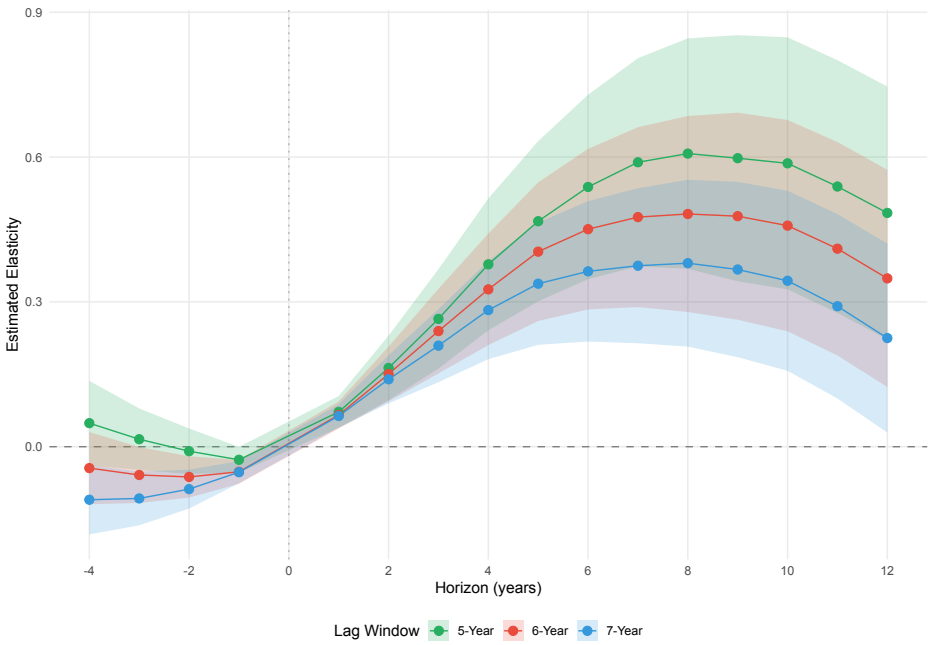


Figure 8: Estimates of Elasticity of Number of PhDs to Change in Government Funding: Robustness for Horizon. Government funding increases are measured as the difference in level of cumulated FM shocks between $t-k$ and t. Weights are measured using the average share of agency R&D obligations to universities, university operated FFRDCs, and not for profits relative to total yearly R&D obligations from the 1999-2019 in the Survey of Federal Funds for Research and Development. Weights: NIH=71%,NSF = 87%,NASA = 13%,DOE = 11%, DOD = 5%. Standard errors are clustered by field.

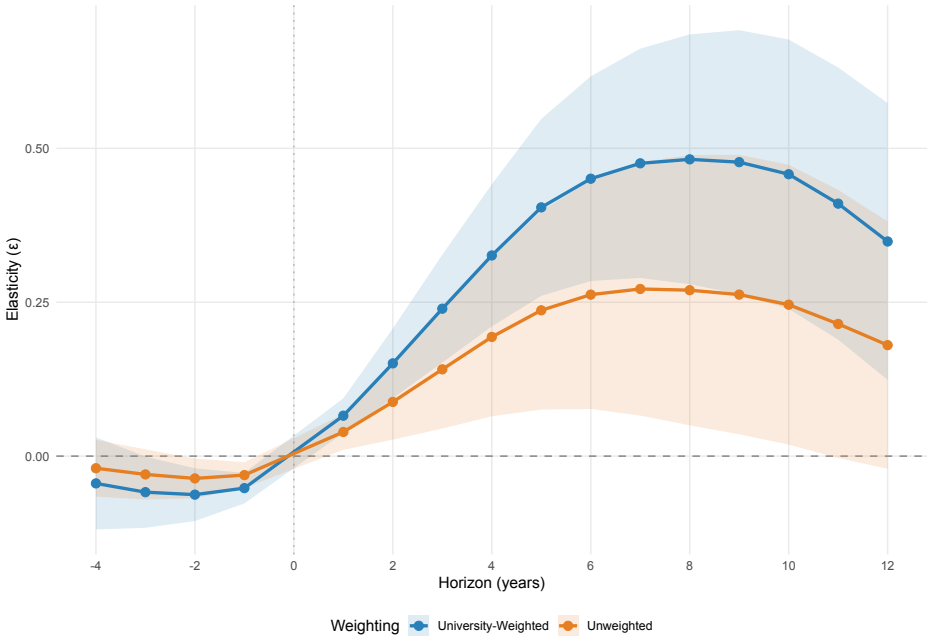


Figure 9: Estimates of Elasticity of Number of PhDs to Change in Government Funding: Robustness to Weighting. Government funding increases are measured as the difference in level of cumulated FM shocks between $t-6$ and t. Weights are measured using the average share of agency R&D obligations to universities, university operated FFRDCs, and not for profits relative to total yearly R&D obligations from the 1999-2019 in the Survey of Federal Funds for Research and Development. Weights: NIH=71%,NSF = 87%,NASA = 13%,DOE = 11%, DOD = 5%. Standard errors are clustered by field.

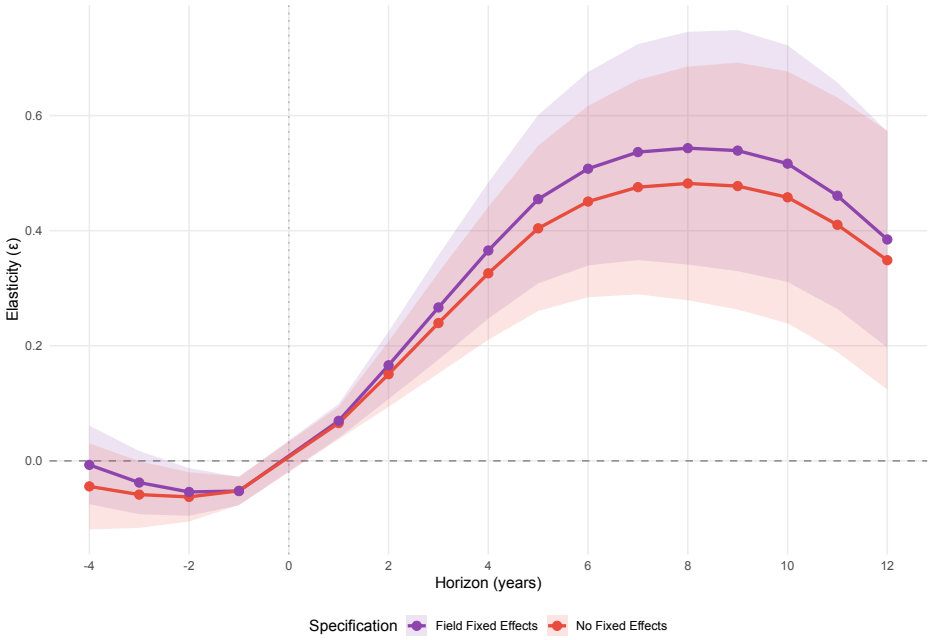


Figure 10: Estimates of Elasticity of Number of PhDs to Change in Government Funding: Robustness to Field FEs. Government funding increases are measured as the difference in level of cumulated FM shocks between $t-6$ and t. Weights are measured using the average share of agency R&D obligations to universities, university operated FFRDCs, and not for profits relative to total yearly R&D obligations from the 1999-2019 in the Survey of Federal Funds for Research and Development. Weights: NIH=71%,NSF = 87%,NASA = 13%,DOE = 11%, DOD = 5%. Field fixed effects allow intercept to vary by field. Standard errors are clustered by field.

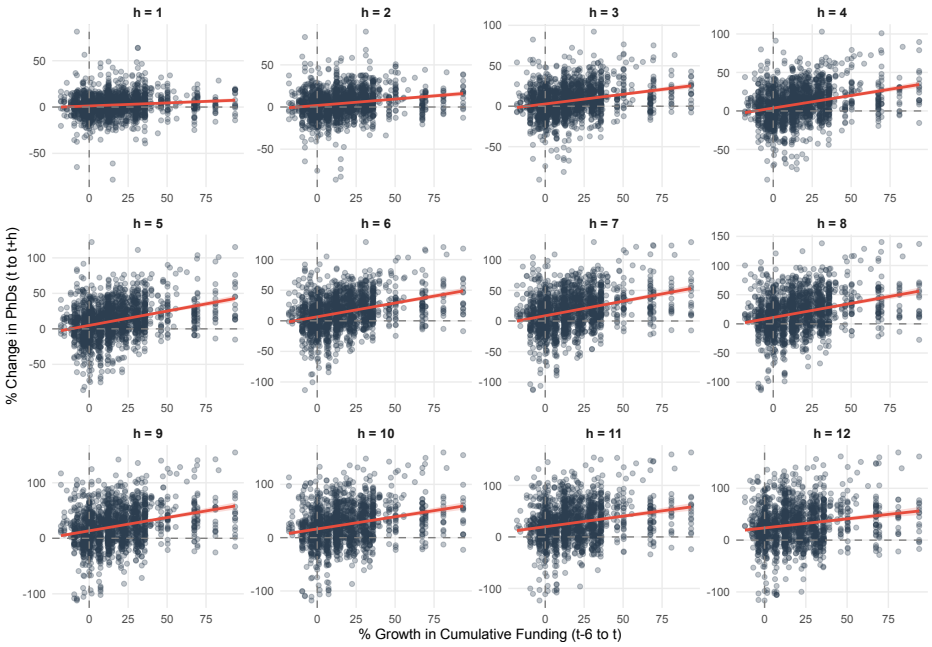


Figure 11: Scatterplots of field growth rates at different horizons. A dot is a field-year. Government funding increases are measured as the difference in level of cumulated FM shocks between $t-6$ and t with t running from 1970 onwards. Weights are measured using the average share of agency R&D obligations to universities, university operated FFRDCs, or not for profits relative to total yearly R&D obligations from the 1999-2019 in the Survey of Federal Funds for Research and Development. Weights: NIH=71%,NSF = 87%,NASA = 13%,DOE = 11%, DOD = 5%.

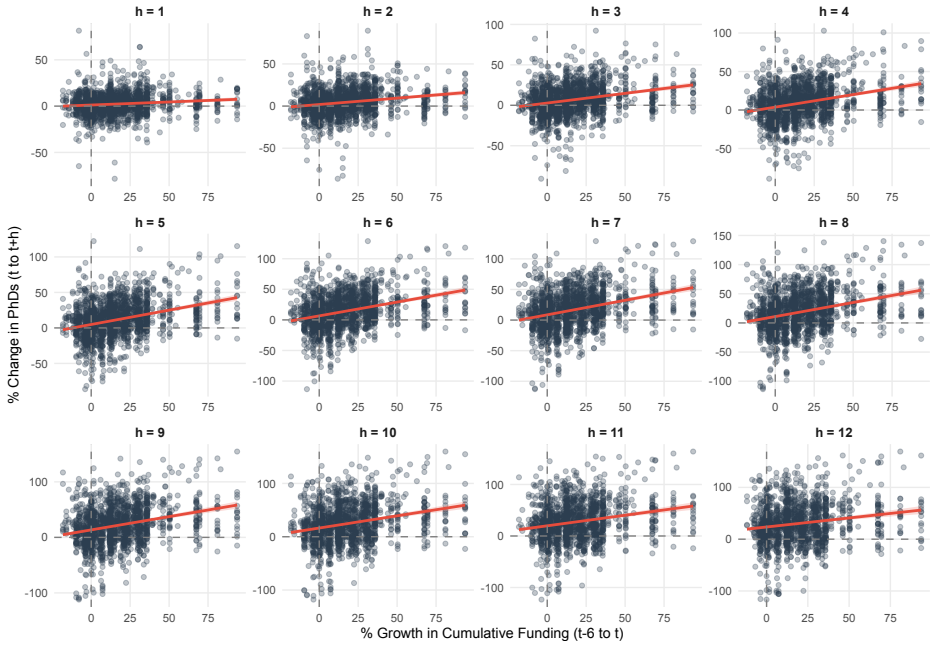


Figure 12: Scatterplots of field growth rates at different horizons. A dot is a field-year. Government funding increases are measured as the difference in level of cumulated FM shocks between $t-6$ and t with t running from 1970 onwards. No weights are used

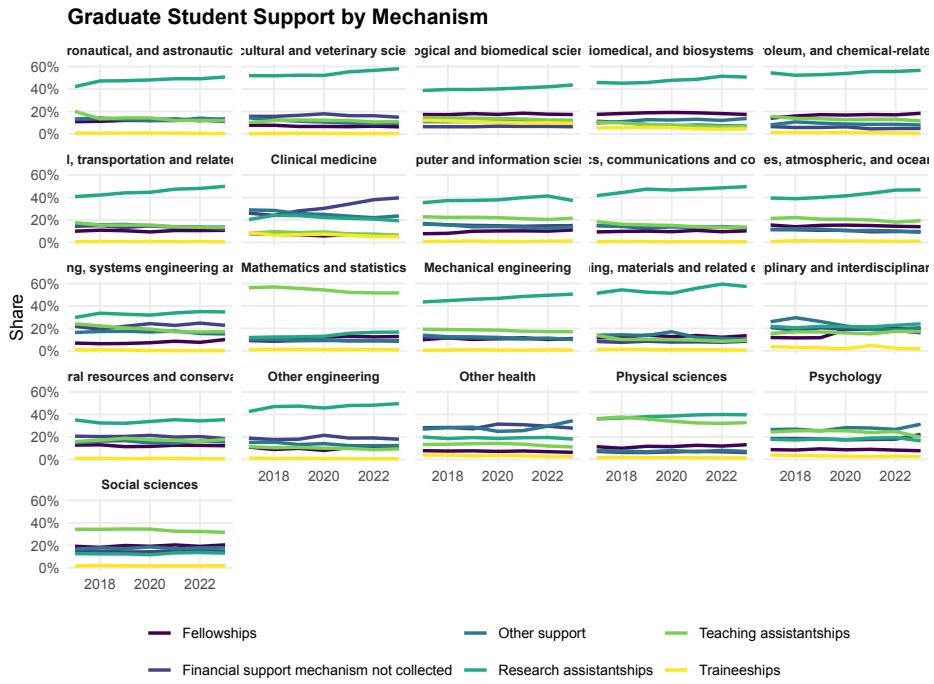


Figure 13: Primary mechanism of support by field from the Survey of Graduate Students and Post-Docs in Science Engineering.

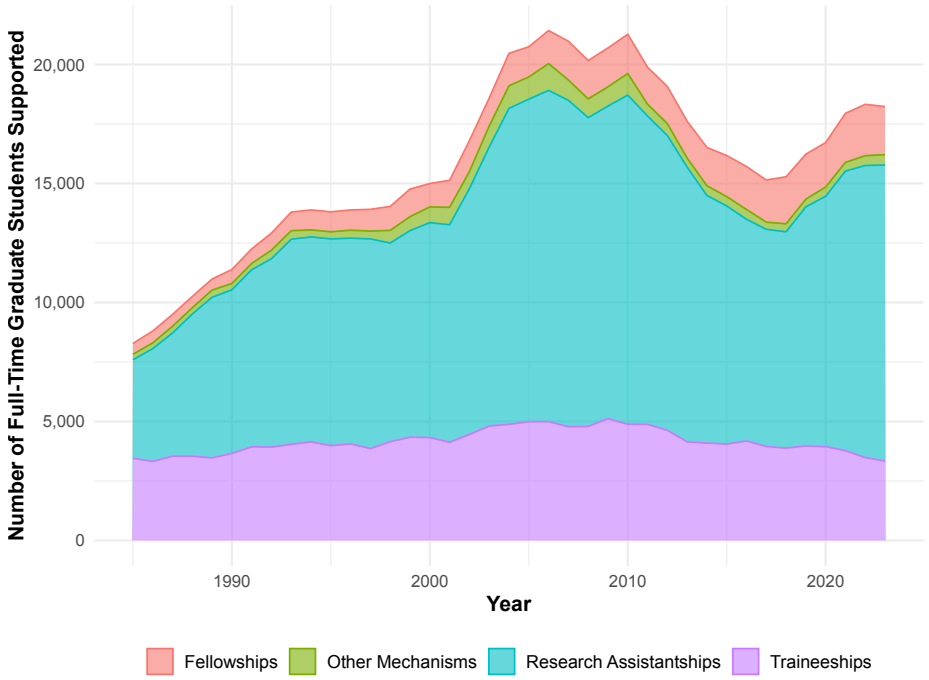


Figure 14: NIH graduate student support from the NIH databook.

B Firm Level Results for Baseline Model

This section contains proofs of results for the firm side of the baseline model.

B.1 The Static Aggregation

The final goods sector is competitive and aggregates sectoral goods together using a Cobb-Douglas production technology. The problem of these firms is given by:

$$\max_{\{Y_s\}} \prod_{s \in S} Y_s^{\beta_s} - \sum_{s \in S} P_s Y_s$$

This implies the standard Cobb-Douglas constant expenditure shares:

$$P_s Y_s = \beta_s Y$$

Each of the sectoral goods aggregators aggregates a unit measure of varieties in Cobb-Douglas manner:

$$\max_{x_s} P_s \exp\left\{ \int_0^1 \ln(y_s(i)) di \right\} - \int_0^1 p_s(i) y_s(i) di$$

This implies that:

$$\forall i \; P_s Y_s = p_s(i) y_s(i)$$

Each intermediate good manufacturer has access to a technology that converts production labor into intermediate goods with labor productivity z_i. As will become clear when I describe the structure of the markets for innovation in this economy, firms that operate in a given product line have a productivity equal to λ times a public technology. Bertrand competition and the assumption of perfect substitution within a product line then imply that the price is set at the marginal cost of the public technology, which then implies a fixed markup equal to λ, the step-size, given the yet to be described innovation structure.

$$y_s(i) = z_s(i) l_s(i)$$

$$\begin{aligned}
\mu_s(i) &= p_s(i) \times \frac{z_s(i)}{w} \\
&= \frac{w}{z_s(-i)} \times \frac{z_s(i)}{w} \\
&= \frac{\lambda w}{z_s(i)} \times \frac{z_s(i)}{w} \\
&= \lambda \\
y_s(i) &= \frac{P_s Y_s}{p_s(i)} \\
z_s(i) l_s(i) &= \frac{\beta_s Y}{\lambda w} z_s(i) \\
l_s(i) &= \frac{\beta_s Y}{\lambda w} \\
\pi_s(i) &= (\mu_s(i) - 1) \frac{w l_i}{z_i} \\
&= \left(\frac{\lambda - 1}{\lambda} \right) \beta_s Y
\end{aligned}$$

What is the value of a firm at any given point in time? Since firms are born and die due to reasons outside their control, their value is at any time is given by the solution to the following rather mechanical HJB:

$$(r_t + \iota_{st}) V_{st} - \dot{V}_{st} = \max_{I_s} \left\{ \pi_s - \epsilon_{st} \tilde{w}_{st} I_s^{\gamma} + I_s V_{st} \right\}$$

Along a BGP, firm value grows at a constant rate g, the return on assets r is constant, and the rate of displacement in a sector s, ι_s is constant. It follows that the value of a firm operating a product line is just:

$$V_s = \frac{\left(\frac{\lambda-1}{\lambda}\right)\beta_s A L^P - \epsilon_s \tilde{w}_{st} \iota_s^\gamma}{r + \iota_s - g}$$

Having characterized the choices of all firms producing physical quantities in the economy, I can now derive a simple aggregate production function. To do so, proceed up from the optimal choices of intermediate goods producers.

$$\begin{aligned}
\ln(Y_s) &= \int_0^1 \ln(y_s(i))di \\
&= \int_0^1 \ln(z_s(i)) + \ln\left(\frac{\beta_s Y}{\lambda w_0}\right)di \\
&= \int_0^1 \ln(z_s(i))di + \ln\left(\frac{\beta_s Y}{\lambda w_0}\right) \\
Y_s &= \underbrace{\exp\left(\int_0^1 \ln(z_s(i))di\right)}_{\equiv Z_s} \frac{\beta_s Y}{\lambda w_0}
\end{aligned}$$

Next, use the production function of the final goods aggregator to derive an expression for the wage of production labor in terms of an aggregate TFP index:

$$\begin{aligned}
Y &= \prod_{s\in S} Y_s^{\beta_s} \\
&= \frac{Y}{\lambda w_0} \prod_{s\in S} (\beta_s Z_s)^{\beta_s} \\
&= \frac{Y}{\lambda w_0} \prod_{s\in S} \exp\left\{\beta_s \ln(\beta_s Z_s)\right\} \\
&= \frac{Y}{\lambda w_0} \exp\left\{\sum_{s\in S} \beta_s \ln(\beta_s Z_s)\right\} \\
&= \frac{Y}{\lambda w_0} \exp\left\{\sum_{s\in S} \beta_s \ln(\beta_s)\right\} \exp\left\{\sum_{s\in S} \beta_s \ln(Z_s)\right\} \\
w_0 &= \frac{1}{\lambda} \exp\left\{\sum_{s\in S} \beta_s \ln(\beta_s)\right\} \exp\left\{\sum_{s\in S} \beta_s \ln(Z_s)\right\}
\end{aligned}$$

Now, return to the problem of intermediate goods producers and observe that the aggregate quantity of unspecialized labor used in production is given by:

$$\begin{aligned} L^P &= \sum_{s\in S}\int_0^1 l_s(i)di \\ &= \sum_{s\in S}\frac{\beta_s Y}{\lambda w_0} \\ Y &= w_0\lambda L^P \\ Y &= \underbrace{\exp\left\{\sum_{s\in S}\beta_s\ln(\beta_s)\right\}}_{\equiv A}\exp\left\{\sum_{s\in S}\beta_s\ln(Z_s)\right\}L^P \end{aligned}$$

Notice that A is a function only of exogenous and time invariant parameters. This implies that the growth rate of output is just given by:

$$g = \ln(\lambda)\sum_{s\in S}\beta_s\iota_s$$

B.2 Innovation Results

Innovations occur in each sector via the actions of a continuum of idea producing firms. Each of these firms have access to a linear technology that takes as input a collection of different types of knowledge and outputs ideas according to the following idea production function:

$$I_s = \frac{1}{\epsilon_s^{\frac{1}{\gamma}}}\left(\prod_{k\in K}H_{s,k}^{\alpha_{s,k}}\right)^{\frac{1}{\gamma}}$$

When an idea producing firm has an idea, they improve the productivity of a random variety in the sector by a factor λ. The existing production technique for that technology diffuses to the public, which ensures that there is a constant (and equal across sectors) price-cost markup equal to λ.

To characterize the problem, I proceed in two steps. In the first step, I derive the cost-minimizing

mix of inputs, while in the second step, I find the profit-maximizing level of output, taking as given the cost-minimizing mix. The cost-minimization problem of an idea producing firm can be written as:

$$\min_{H_{s,k}} \sum_{k \in K} w_k H_{s,k}$$

Subject to:

$$\bar{I} \leq \frac{1}{\epsilon_s^{\frac{1}{\gamma}}} \left(\prod_{k \in K} H_{s,k}^{\alpha_{s,k}} \right)^{\frac{1}{\gamma}}$$

This has first order condition:

$$w_k = (-\mu) \frac{\alpha_{s,k}}{\gamma \epsilon_s^{\frac{1}{\gamma}}} \frac{\left(\prod_{k \in K} H_{s,k}^{\alpha_{s,k}} \right)^{\frac{1}{\gamma}}}{H_{s,k}}$$

$$
\begin{aligned}
H_{s,k} &= (-\mu)\frac{\alpha_{s,k}}{\gamma\epsilon_s^{\gamma}}\frac{\left(\prod_{k\in K} H_{s,k}^{\alpha_{s,k}}\right)^{\frac{1}{\gamma}}}{w_k} \\
H_{s,k'} &= \frac{\alpha_{s,k'}}{\alpha_{s,k}}\frac{w_k}{w_{k'}}H_{s,k} \\
&= \frac{\alpha_{s,k'}}{w_{k'}}\frac{w_k H_{s,k}}{\alpha_{s,k}} \\
\prod_{k'\in K} H_{s,k'}^{\alpha_{s,k'}} &= \frac{w_k H_{s,k}}{\alpha_{s,k}}\prod_{k'\in K}\left(\frac{\alpha_{s,k'}}{w_{k'}}\right)^{\alpha_{s,k'}} \\
I_s &= \frac{1}{\epsilon_s^{\frac{1}{\gamma}}}\left(\frac{w_k H_{s,k}}{\alpha_{s,k}}\right)^{\frac{1}{\gamma}}\left(\prod_{k'\in K}\left(\frac{\alpha_{s,k'}}{w_{k'}}\right)^{\alpha_{s,k'}}\right)^{\frac{1}{\gamma}} \\
w_k H_{s,k} &= \epsilon_s\left(\prod_{k'\in K}\left(\frac{w_{k'}}{\alpha_{s,k'}}\right)^{\alpha_{s,k'}}\right)\alpha_{s,k} I_s^{\gamma} \\
C(I_s) &= \epsilon_s\left(\prod_{k\in K}\left(\frac{w_k}{\alpha_{s,k}}\right)^{\alpha_{s,k}}\right) I_s^{\gamma}
\end{aligned}
$$

From here, we must now derive the optimal choice of I_s in this economy. To that end, the problem of the firm after selecting a cost-minimizing mix of inputs is to solve:

$$
\max_{I_s} V_s I_s - \epsilon_s\left(\prod_{k\in K}\left(\frac{w_k}{\alpha_{s,k}}\right)^{\alpha_{s,k}}\right) I_s^{\gamma}
$$

Let $\tilde{w}_s \equiv \left(\prod_{k\in K}\left(\frac{w_k}{\alpha_{s,k}}\right)^{\alpha_{s,k}}\right)$

This implies that:

$$
\begin{aligned}
V_s &= \gamma\epsilon_s\tilde{w}_s I_s^{\gamma-1} \\
I_s &= \left[\frac{\gamma\epsilon_s\tilde{w}_s}{V_s}\right]^{\frac{1}{1-\gamma}}
\end{aligned}
$$

Using the primal approach and characterizing profit maximization also provides one with a

more convenient way to write the equilibrium pricing condition the mutual fund. Specfically note that:

$$\max \frac{V_s}{\epsilon_s^{\frac{1}{\gamma}}} \left(\prod_{k \in K} H_{s,k}^{\alpha_{s,k}} \right)^{\frac{1}{\gamma}} - \sum_k w_k H_{s,k}$$

This implies that:

$$\frac{V_s \alpha_{s,k} I_s}{\gamma} = w_k H_{s,k}$$

Summing across k then implies that:

$$\sum_k w_k H_{s,k} = \frac{V_s I_s}{\gamma}$$

This implies that we can write the differential equation that governs the value of the aggregate mutual fund as:

$$\frac{dV^{MF}}{dt} = (r_t - g_t) V_t^{MF} - \left(\Pi_t - \frac{1}{\gamma} \sum_s V_{st} I_{st} \right)$$

C Worker Problems

In this section I solve the baseline model of occupational choice in closed form. Suppose now that there exists a balanced growth path with growth rate g. As is standard, begin by defining detrended variables:

$$\tilde{a} \equiv ae^{-gt}$$

$$\tilde{c} \equiv ce^{-gt}$$

$$w_{t,k} \equiv w_k e^{gt}$$

Note that:

$$\dot{a} = \frac{d}{dt}(\tilde{a}e^{gt}) = e^{gt}(\dot{\tilde{a}} + g\tilde{a})$$

Use this to write the budget constraint

$$\dot{a} = (r+\delta)\tilde{a}e^{gt} + w_k e^{gt} z_k(\tau) - \tilde{c}e^{gt}$$

as

$$\dot{\tilde{a}} = (r+\delta-g)\tilde{a} + w_k z_k(\tau) - \tilde{c}$$

Now guess $V_t(a, z_0, \tau) = e^{(1-\sigma)gt}\tilde{V}(\tilde{a}, z_0, \tau)$

It follows that:

$$\frac{\partial V_t(a, z_0, \tau)}{\partial a} = e^{-\sigma t}\frac{\partial \tilde{V}(\tilde{a}, z_0, \tau)}{\partial \tilde{a}}$$

Similarly:

$$\frac{\partial V}{\partial \tau} = (1-\sigma)ge^{(1-\sigma)t}\tilde{V} + e^{(1-\sigma)gt}\frac{\partial \tilde{V}}{\partial \tau}\frac{d\tilde{a}}{dt} = (1-\sigma)ge^{(1-\sigma)t}\tilde{V} - e^{(1-\sigma)gt}\frac{\partial \tilde{V}}{\partial \tau}g\tilde{a}$$

$$\frac{\partial \tilde{V}}{\partial \tau} = e^{-(1-\sigma)gt}\frac{\partial V}{\partial \tau}$$

Together, these expressions imply a HJB given by:

$$(\rho + \delta + (\sigma-1)g)\tilde{V} = \max_{\tilde{c}} \left\{ \frac{\tilde{c}^{1-\sigma}}{1-\sigma} + \frac{\partial \tilde{V}}{\partial \tilde{a}}[(r+\delta-g)\tilde{a} + w_k z_0 h(\tau) - \tilde{c}] + \frac{\partial \tilde{V}}{\partial \tau} \right\} \tag{40}$$

C.1 Solving the detrended problem

Having now constructed a normalized value function, I now characterize the solution to the normalized value function. To do so, suppose that an individual is born with an endowment of skills that grow according to a deterministic trend. Let τ denote the time of birth of an individual. Each field has a growth profile given by $h(\tau)$, where I will specify $h(\tau)$ later as necessary. To characterize the worker problem, I proceed through a strategy of guess and verify. The state variables of this problem are a: the asset position of agents, z_0: the draw of the agents, τ the agent's age, and k their choice of field at time of birth.

The HJB for this problem is given by:

$$(\rho + \delta + (\sigma-1)g)V(a, z_0, \tau; k) = \max_{c} \left\{ \frac{c^{1-\sigma}}{1-\sigma} + \frac{\partial V}{\partial \tau} + \frac{\partial V}{\partial a}\left[(r+\delta-g)a + w_k z_0\big[1 + (\kappa_k - 1)(1 - \exp\{-\lambda_k \tau\})\big] - c\right]\right\} \tag{41}$$

Guess that the value function has the form

$$V(a, z_0, \tau) = \Gamma \frac{(a + \psi(\tau) z_0)^{1-\sigma}}{1-\sigma}$$

This implies the following set of first order conditions:

$$\frac{\partial V}{\partial a} = \Gamma (a + \psi(\tau) z_0)^{-\sigma} \tag{42}$$

$$\frac{\partial V}{\partial \tau} = \Gamma (a + \psi(\tau) z_0)^{-\sigma} \dot{\psi}(\tau) z_0 \tag{43}$$

The first order condition of the HJB requires that:

$$\Gamma (a + \psi(\tau) z_0)^{-\sigma} = c^{-\sigma} \iff c = \Gamma^{-\frac{1}{\sigma}} (a + \psi(\tau) z_0) \tag{44}$$

Plugging our guess into the HJB, we have that:

$$\begin{aligned}(\rho + \delta + (\sigma - 1) g) \Gamma \frac{(a + \psi(\tau) z_0)^{1-\sigma}}{1-\sigma} - \Gamma^{\frac{\sigma-1}{\sigma}} \frac{(a + \psi(\tau) z_0)^{1-\sigma}}{1-\sigma} \\ = \Gamma (a + \psi(\tau) z_0)^{-\sigma} [(r + \delta - g) a + w_k z_0 h(\tau) - \Gamma^{-\frac{1}{\sigma}} (a + \psi(\tau) z_0)] \\ + \Gamma (a + \psi(\tau) z_0)^{-\sigma} \dot{\psi}(\tau) z_0\end{aligned}$$

Dividing through by $\Gamma (a + \psi(\tau) z_0)^{-\sigma}$, it follows:

$$\begin{aligned}(a + \psi(\tau) z_0) \left[\frac{(\rho + \delta + (\sigma - 1) g) - \sigma \Gamma^{-\frac{1}{\sigma}}}{1-\sigma} \right] = (r + \delta - g) a + w_k z_0 h(\tau) \\ + \dot{\psi}(\tau) z_0\end{aligned}$$

Matching coefficients for a, we have that:

$$\begin{aligned}
\rho + \delta + (\sigma - 1)g - \sigma\Gamma^{-\frac{1}{\sigma}} &= (1 - \sigma)(r + \delta - g) \\
\sigma\Gamma^{-\frac{1}{\sigma}} &= \rho + \delta + (\sigma - 1)g + (\sigma - 1)(r + \delta) - (\sigma - 1)g \\
\sigma\Gamma^{-\frac{1}{\sigma}} &= \rho + \delta + (\sigma - 1)(r + \delta) \\
\Gamma^{-\frac{1}{\sigma}} &= \frac{\rho + \delta + (\sigma - 1)(r + \delta)}{\sigma} \\
\Gamma &= \left[\frac{\rho + \delta + (\sigma - 1)(r + \delta)}{\sigma}\right]^{-\sigma}
\end{aligned}$$

Matching coefficients for z_0, and using the value of Γ derived above we have that:

$$\begin{aligned}
\psi(\tau)\big[r + \delta - g\big] &= w_k h(\tau) + \frac{d\psi}{d\tau} \\
\frac{d\psi}{d\tau} &= \big[r + \delta - g\big]\psi(\tau) - w_k h_k(\tau)
\end{aligned}$$

Although one could solve this in general under sufficient conditions on $h(\tau)$, here I will specify a functional form for $h_k(\tau)$.

Rewrite this as a differential equation:

$$\dot{\psi}(\tau) = (r + \delta - g)\psi(\tau) - w_k \underbrace{\big[1 + (\kappa_k - 1)(1 - e^{-\lambda_k \tau})\big]}_{= h_k(\tau)}$$

This is a canonical first order differential equation that can be solved via the method of integrating factors. The integrating factor is $e^{-(r+\delta-g)t}$.

$$
\begin{aligned}
e^{-(r+\delta-g)t}\frac{d\psi}{d\tau} - (r+\delta-g)e^{-(r+\delta-g)\tau}\psi(\tau) &= -w_k e^{-(r+\delta-g)\tau}[1+(\kappa_k-1)(1-e^{-\lambda_k\tau})] \\
\frac{d}{dt}\left[e^{-(r+\delta-g)t}\psi(\tau)\right] &= -w_k e^{-(r+\delta-g)\tau}[1+(\kappa_k-1)(1-e^{-\lambda_k t})] \\
&= -w_k e^{-(r+\delta-g)\tau} - w_k e^{-(r+\delta-g)\tau}(\kappa_k-1)+ \\
& w_k(\kappa_k-1)e^{-(r+\delta-g+\lambda_k)\tau} \\
&= -\kappa_k w_k e^{-(r+\delta-g)\tau} + w_k(\kappa_k-1)e^{-(r+\delta-g+\lambda_k)\tau}
\end{aligned}
$$

Integrating both sides, we recover that:

$$
\begin{aligned}
e^{-(r+\delta-g)\tau}\psi(\tau) &= \int_0^\tau -\kappa_k w_k e^{-(r+\delta-g)s} + w_k(\kappa_k-1)e^{-(r+\delta-g+\lambda_k)s}ds \\
&= \frac{\kappa_k w_k}{r+\delta-g}e^{-(r+\delta-g)s}\Big|_0^\tau + \frac{w_k(1-\kappa_k)}{r+\delta-g+\lambda_k}e^{-(r+\delta-g+\lambda_k)s}\Big|_0^\tau + C \\
&= \frac{\kappa_k w_k}{r+\delta-g}[e^{-(r+\delta-g)\tau}-1] - \frac{w_k(\kappa_k-1)}{r+\delta-g+\lambda_k}[e^{-(r+\delta-g+\lambda_k)\tau}-1] + C \\
\psi(\tau) &= \frac{\kappa_k w_k}{r+\delta-g}[1-e^{(r+\delta-g)\tau}] - \frac{w_k(\kappa_k-1)}{r+\delta-g+\lambda_k}[e^{-\lambda_k\tau}-e^{(r+\delta-g)\tau}] + Ce^{(r+\delta-g)\tau}
\end{aligned}
$$

Assume that $\lim_{t\to\infty}\psi(\tau)$ exists. Then it must be the case that C exactly isolates the explosive terms, so we have that:

$$
\begin{aligned}
\psi(\tau) &= \frac{\kappa_k w_k}{r+\delta-g} - \frac{w_k(\kappa_k-1)e^{-\lambda_k\tau}}{r+\delta-g+\lambda_k} \\
&= w_k\left[\frac{\kappa_k}{r+\delta-g} - \frac{(\kappa_k-1)e^{-\lambda_k\tau}}{r+\delta-g+\lambda_k}\right]
\end{aligned}
$$

What remains is to verify that $\psi(\tau)$ satisfies the differential equation. To that end, note that:

Do this later

This verifies the conjecture. A couple other things to note that are helpful is that:

Let $x(\tau) \equiv a + \psi(\tau)z_0$.

$$\dot{x}(\tau) \equiv \dot{a} + \dot{\psi}(\tau)z_0$$

$$\begin{aligned}
\dot{a} &= (r + \delta - g)a + w_k h(\tau)z_0 - c \\
\dot{x}(\tau) - \dot{\psi}(\tau)z_0 &= (r + \delta - g)(x - \psi(\tau)z_0) + w_k h(\tau)z_0 - c \\
\dot{x}(\tau) &= (r + \delta - g)x(\tau) + \dot{\psi}(\tau)z_0 - \big[\underbrace{(r + \delta - g)\psi(\tau)z_0 - w_k h(\tau)z_0}_{=\dot{\psi}(\tau)z_0}\big] - c \\
\dot{x}(\tau) &= (r + \delta - g)x(\tau) - c
\end{aligned}$$

Hence:

$$\begin{aligned}
\dot{x}(\tau) &= \underbrace{(r + \delta - g - \Gamma^{-\frac{1}{\sigma}})}_{\equiv\tilde{\mu}} x(\tau) \\
\dot{a} &= \tilde{\mu}a + \tilde{\mu}\psi(\tau)z_0 - \dot{\psi}(\tau)z_0 \\
\dot{a} &= \tilde{\mu}a + \big[\tilde{\mu}\psi(\tau) - \dot{\psi}(\tau)\big]z_0 \\
\dot{a} - \tilde{\mu}a &= \big[\tilde{\mu}\psi(\tau) - \dot{\psi}(\tau)\big]z_0
\end{aligned}$$

We want to solve this forward for $a(\tau)$. Neglecting arguments as a function of age, consider the ansatz $a = b - \psi z$ for some function b. This implies that $\dot{a} = \dot{b} - \dot{\psi}z$

Then we have that:

$$\dot{b} - \dot{\psi}z - \tilde{\mu}(b - \psi z) = \mu\psi z - \dot{\psi}z$$
$$\dot{b} - \tilde{\mu}b - \dot{\psi}z + \tilde{\mu}\psi z = \mu\psi z - \dot{\psi}z$$
$$\dot{b} - \tilde{\mu}b = 0$$
$$\dot{b} = \tilde{\mu}b$$
$$\frac{d}{dt}[\ln(b)] = \tilde{\mu}$$
$$\ln(b) = \tilde{\mu}\tau + C$$
$$b = \tilde{C}e^{\tilde{\mu}\tau}$$
$$a(\tau) + \psi(\tau)z = \tilde{C}e^{\tilde{\mu}\tau}$$

This is the general solution, to identify the particular solution we use the initial condition that $a(0) = 0$ to identify that:

$$\psi(0)z = \tilde{C}$$

Hence the particular solution is given by:

$$a(\tau) = z_0\psi(0)e^{\tilde{\mu}\tau} - z_0\psi(\tau) \tag{45}$$

Notice that this is expressed in terms of detrended variables. Given an initial distribution wealth - which I will take to be zero - we can then recover aggregate asset holdings by type k by integrating this function over the age distribution.

Hence the value function has form given by:

$$V(a, z_0, \tau) = \left[\frac{(\rho + \delta) + (\sigma - 1)(r + \delta)}{\sigma}\right]^{-\sigma} \frac{\left(a + w_k z_0 \left[\frac{\kappa_k}{r+\delta-g} - \frac{(\kappa_k - 1)e^{-\lambda_k \tau}}{r+\delta-g+\lambda_k}\right]\right)^{1-\sigma}}{1 - \sigma}$$

The (detrended) policy function for consumption is given by:

$$c = \left[\frac{\rho + \delta + (\sigma - 1)(r + \delta)}{\sigma}\right]\left(a + w_k z_0 \left[\frac{\kappa_k}{r + \delta - g} - \frac{(\kappa_k - 1)e^{-\lambda_k \tau}}{r + \delta - g + \lambda_k}\right]\right)$$

Assuming that individuals start with 0 financial wealth, the value at birth is given by:

$$V(0, z_0, 0, k) \propto \frac{\left(w_k z_0 \left[\frac{r+\delta-g+\kappa_k\lambda_k}{r+\delta-g+\lambda_k}\right]\right)^{1-\sigma}}{1-\sigma}$$

This immediately implies that the problem of a new-born along the balanced growth path can be expressed as a static problem:

$$k^* = \arg\max_{k \in K} \left\{ w_k \underbrace{\left[\frac{r + \delta - g + \kappa_k \lambda_k}{r + \delta - g + \lambda_k}\right]}_{\equiv \Psi_k} z_k \right\} \tag{46}$$

Because the problem takes this form and the skill distribution is distributed multivariate Frechet, we can characterize choice probabilities directly using results from extreme value theory. Specifically, one can analytically derive the labor supply of a generation at age 0 and then use that to characterize aggregate labor supply:

$$\mathbb{P}_k = \sum_f \left\{ \underbrace{\frac{\left[\sum_{k'} (\omega_{fk'} T_{k'} (\Psi_{k'} w_{k'})^\theta)^{\frac{1}{1-\rho_f}}\right]^{1-\rho_f}}{\sum_f \left[\sum_{k'} (\omega_{fk'} T_{k'} (\Psi_{k'} w_{k'})^\theta)^{\frac{1}{1-\rho_f}}\right]^{1-\rho_f}}}_{\equiv \mathbb{P}^f} \times \underbrace{\frac{(\omega_{fk} T_k (\Psi_k w_k)^\theta)^{\frac{1}{1-\rho_f}}}{\sum_{k'} (\omega_{fk'} T_{k'} (\Psi_{k'} w_{k'})^\theta)^{\frac{1}{1-\rho_f}}}}_{\equiv \mathbb{P}_{k|f}} \right\}$$

$$\mathcal{L}_k(0) = T_k^{\frac{1}{\theta}} \mathbb{P}_k^{1-\frac{1}{\theta}} \Gamma\left(1 - \frac{1}{\theta}\right) \times \left[\sum_f \left\{ \left[\sum_{k'} (\omega_{fk'} T_{k'} w_{k'}^\theta)^{\frac{1}{1-\rho_f}}\right]^{1-\rho_f} \times \frac{(\omega_{fk} T_k w_k^\theta)^{\frac{1}{1-\rho_f} - 1} \omega_{fk}}{\sum_{k'} (\omega_{fk'} T_{k'} w_{k'}^\theta)^{\frac{1}{1-\rho_f}}} \right\}\right]^{\frac{1}{\theta}}$$

$$\begin{aligned}\mathcal{L}_k(0) &= \frac{\mathbb{P}_k}{w_k\Psi_k}\mathbb{E}[\max_j w_j\Psi_j z_j] \\ &= T_k^{\frac{1}{\theta}}\mathbb{P}_k^{1-\frac{1}{\theta}}\Gamma\left(1-\frac{1}{\theta}\right)\left[\sum_f\left\{\left[\sum_{k'}(\omega_{fk'}T_{k'}(\Psi_{k'}w_{k'})^\theta)^{\frac{1}{1-\rho_f}}\right]^{1-\rho_f}\frac{(\omega_{fk}T_k(\Psi_k w_k)^\theta)^{\frac{1}{1-\rho_f}-1}\omega_{fk}}{\sum_{k'}(\omega_{fk'}T_{k'}(\Psi_{k'}w_{k'})^\theta)^{\frac{1}{1-\rho_f}}}\right\}\right]^{\frac{1}{\theta}}\end{aligned}$$

Note again that this is the labor supply of an individual generation *at birth*, when $h(\tau) = 1$. To recover aggregate labor supply, we have to integrate over the distribution of individuals alive at any given point in time. Specifically, and neglecting the time subscript, we have that:

$$\mathcal{L}_k = \int_0^\infty \mathcal{L}_k(\tau)h_k(\tau)\delta\exp\big\{-\delta\tau\big\}d\tau$$

Along a balanced growth path, relative prices and growth rates are constant, which implies that the labor supply of any two generations are identical. The consequence of this is that we can write the above as:

$$\begin{aligned}\mathcal{L}_k &= \delta\mathcal{L}_k(0)\int_0^\infty h_k(\tau)\exp\{-\delta\tau\}d\tau \\ &= \delta\mathcal{L}_k(0)\int_0^\infty \big[1+(\kappa_k-1)(1-\exp\{-\lambda_k\tau\}\big]\exp\{-\delta\tau\}d\tau \\ &= \delta\mathcal{L}_k(0)\int_0^\infty \big[\exp\{-\lambda_k\tau\}+\kappa_k-\kappa_k\exp\{-\lambda_k\tau\}\big]\exp\{-\delta\tau\}d\tau \\ &= \delta\mathcal{L}_k(0)\int_0^\infty \exp\{-(\lambda_k+\delta)\tau\}+\kappa_k\exp\{-\delta\tau\}-\kappa_k\exp\{-(\lambda_k+\delta)\tau\}d\tau \\ &= \delta\mathcal{L}_k(0)\left[\frac{\kappa_k}{\delta}+\frac{(1-\kappa_k)}{\delta+\lambda_k}\right] \\ &= \mathcal{L}_k(0)\left[\frac{\delta+\kappa_k\lambda_k}{\delta+\lambda_k}\right]\end{aligned}$$

This delivers balanced growth path labor *supply* as a function of primitives and equilibrium

relative wages, the aggregate growth rate, and interest rate. To complete the characterization of the household side of the model, we must also derive the aggregate Euler equation, which is what will generate the equilibrium interest rate. Starting from the consumption optimal consumption choice, we have that:

$$
\begin{aligned}
c &= \left[\frac{\rho+\delta+(\sigma-1)(r+\delta)}{\sigma}\right](a+\psi(\tau)z_0)\\
\dot{c} &= \left[\frac{\rho+\delta+(\sigma-1)(r+\delta)}{\sigma}\right](\dot{a}+\dot{\psi}(\tau)z_0)\\
\dot{c} &= \left[\frac{\rho+\delta+(\sigma-1)(r+\delta)}{\sigma}\right]\left(\big(r+\delta-g\big)a+w_k z_0 h_k(\tau)-c+\dot{\psi}(\tau)z_0\right)\\
\dot{c} &= \left[\frac{\rho+\delta+(\sigma-1)(r+\delta)}{\sigma}\right]\left(\big(r+\delta-g\big)a+w_k z_0 h_k(\tau)-c+\big(r+\delta-g\big)\psi(\tau)z_0-w_k z_0 h_k(\tau)\right)\\
\dot{c} &= \left[\frac{\rho+\delta+(\sigma-1)(r+\delta)}{\sigma}\right]\left(\big(r+\delta-g\big)a-c+\big(r+\delta-g\big)\psi(\tau)z_0\right)\\
\dot{c} &= c\left(\big(r+\delta-g\big)-\Gamma^{-\frac{1}{\sigma}}\right)\\
\frac{\dot{c}}{c} &= \big(r+\delta-g\big)-\left[\frac{\rho+\delta+(\sigma-1)(r+\delta)}{\sigma}\right]\\
\frac{\dot{c}}{c} &= \frac{r-\rho-\sigma g}{\sigma}
\end{aligned}
$$

Note that this is the Euler equation in *normalized consumption*. To recover the Euler equation in levels, observe that since

$$
\begin{aligned}
\tilde{c} &= ce^{gt}\\
\dot{\tilde{c}} &= \dot{c}e^{gt}+gce^{gt}\\
\dot{\tilde{c}} &= ce^{gt}\left[\frac{\dot{c}}{c}+g\right]\\
\frac{\dot{\tilde{c}}}{\tilde{c}} &= \frac{r-\rho}{\sigma}
\end{aligned}
$$

To derive the aggregate Euler in the normalized problem, observe that we can write aggregate

consumption as:

$$C = \sum_{k \in K} \mathbb{P}_k \int_0^\infty \mathbb{E}[c(a, z_0, \tau; k)| \text{ choose } k] \delta \exp\big\{ - \delta \tau \big\} d\tau$$

Taking time derivatives, we have that:

$$\begin{aligned}
\dot{C} &= \sum_{k \in K} \mathbb{P}_k \Bigg[\int_0^\infty \mathbb{E}[\dot{c}(a, z_0, \tau; k)| \text{ choose } k] \delta \exp\big\{ - \delta \tau \big\} d\tau \\
&\quad - \delta \int_0^\infty \mathbb{E}[c(a, z_0, \tau; k)| \text{ choose } k] \delta \exp\big\{ - \delta \tau \big\} d\tau + \delta \mathbb{E}[c(0, z_0, 0; k)|\text{choose } k] \Bigg] \\
&= \sum_{k \in K} \mathbb{P}_k \Bigg[\int_0^\infty \left(\frac{r - \rho - \sigma g}{\sigma} \right) \mathbb{E}[c(a, z_0, \tau; k)| \text{ choose } k] \delta \exp\big\{ - \delta \tau \big\} d\tau \\
&\quad - \delta \int_0^\infty \mathbb{E}[c(a, z_0, \tau; k)| \text{ choose } k] \delta \exp\big\{ - \delta \tau \big\} d\tau + \delta \mathbb{E}[c(0, z_0, 0; k)|\text{choose } k] \Bigg] \\
&= \sum_{k \in K} \mathbb{P}_k \Bigg[\left(\frac{r - \rho - \sigma g}{\sigma} \right) C_k - \delta C_k + \delta C_k(0) \Bigg] \\
&= \sum_{k \in K} \mathbb{P}_k C_k \left[\left(\frac{r - \rho - \sigma g}{\sigma} \right) - \delta \left[\frac{C_k - C_k(0)}{C_k} \right] \right] \\
\frac{\dot{C}}{C} &= \sum_{k \in K} \frac{\mathbb{P}_k C_k}{C} \left[\left(\frac{r - \rho - \sigma g}{\sigma} \right) - \delta \left[\frac{C_k - C_k(0)}{C_k} \right] \right] \\
\frac{\dot{C}}{C} &= \left(\frac{r - \rho - \sigma g}{\sigma} \right) - \delta \sum_{k \in K} \frac{\mathbb{P}_k C_k}{C} \left[\frac{C_k - C_k(0)}{C_k} \right]
\end{aligned}$$

We now must derive the expressions for C_k in closed form. To that end, note that:

$$C_k \equiv \Gamma^{-\frac{1}{\sigma}} \int_0^\infty \left(\mathbb{E}[a|k, \tau] + \psi_k(\tau) \mathbb{E}[z_k|k] \right) \delta \exp\{-\delta \tau\} d\tau$$

As asset market clearing will imply that the share weighted sum of average (detrended) asset holdings are constant and equal to the total value of all detrended firms, focus only on the second

portion of the summation.

$$\begin{aligned}
\delta\mathbb{E}[z_k|k]\int_0^\infty \psi_k(\tau)\exp\{-\delta\tau\}d\tau &= \delta w_k\mathbb{E}[z_k|k]\int_0^\infty\left[\frac{\kappa_k}{r+\delta-g}-\frac{(\kappa_k-1)e^{-\lambda_k\tau}}{r+\delta-g+\lambda_k}\right]\exp\{-\delta\tau\}d\tau\\
&= \delta w_k\mathbb{E}[z_k|k]\left[\frac{\kappa_k}{\delta(r+\delta-g)}-\frac{\kappa_k-1}{(r+\delta-g+\lambda_k)(\lambda_k+\delta)}\right]\\
&= \frac{\delta}{\Psi_K}w_k\Psi_k\mathbb{E}[z_k|k]\left[\frac{\kappa_k}{\delta(r+\delta-g)}-\frac{\kappa_k-1}{(r+\delta-g+\lambda_k)(\lambda_k+\delta)}\right]\\
&= \frac{\delta}{\Psi_K}\mathbb{E}[w_k\Psi_k z_k|k]\left[\frac{\kappa_k}{\delta(r+\delta-g)}-\frac{\kappa_k-1}{(r+\delta-g+\lambda_k)(\lambda_k+\delta)}\right]\\
&= \mathbb{E}[w_k\Psi_k z_k|k]\times\frac{\delta}{\Psi_k}\times\left[\frac{\kappa_k}{\delta(r+\delta-g)}-\frac{\kappa_k-1}{(r+\delta-g+\lambda_k)(\lambda_k+\delta)}\right]
\end{aligned}$$

Operating on the last part of the equation:

$$\begin{aligned}
\frac{\kappa_k}{\delta(r+\delta-g)}-\frac{\kappa_k-1}{(r+\delta-g+\lambda_k)(\lambda_k+\delta)} &= \frac{\kappa_k(r+\delta-g+\lambda_k)(\lambda_k+\delta)-\kappa_k\delta(r+\delta-g)+\delta(r+\delta-g)}{\delta(r+\delta-g)(r+\delta-g+\lambda_k)(\lambda_k+\delta)}\\
&= \frac{\kappa_k\delta(r+\delta-g)+\lambda_k\kappa_k(r+\delta-g)+\kappa_k\lambda_k(\lambda_k+\delta)-\kappa_k\delta(r+\delta-g)+\delta(r+\delta-g)}{\delta(r+\delta-g)(r+\delta-g+\lambda_k)(\lambda_k+\delta)}\\
&= \frac{\lambda_k\kappa_k(r+\delta-g)+\kappa_k\lambda_k(\lambda_k+\delta)+\delta(r+\delta-g)}{\delta(r+\delta-g)(r+\delta-g+\lambda_k)(\lambda_k+\delta)}\\
&= \frac{\kappa_k\lambda_k(\lambda_k+\delta)+(\lambda_k\kappa_k+\delta)(r+\delta-g)}{\delta(r+\delta-g)(r+\delta-g+\lambda_k)(\lambda_k+\delta)}\\
&= \frac{1}{\delta(r+\delta-g+\lambda_k)}\left[\frac{\kappa_k\lambda_k}{r+\delta-g}+\frac{\kappa_k\lambda_k+\delta}{\lambda_k+\delta}\right]\\
&= \frac{\Psi_k}{\delta(r+\delta-g+\lambda_k\kappa_k)}\left[\frac{\kappa_k\lambda_k}{r+\delta-g}+\frac{\kappa_k\lambda_k+\delta}{\lambda_k+\delta}\right]
\end{aligned}$$

Hence:

$$\delta\mathbb{E}[z_k|k]\int_0^\infty \psi_k(\tau)\exp\{-\delta\tau\}d\tau = \mathbb{E}[w_k\Psi_k z_k|k]\frac{1}{(r+\delta-g+\lambda_k\kappa_k)}\left[\frac{\kappa_k\lambda_k}{r+\delta-g}+\frac{\kappa_k\lambda_k+\delta}{\lambda_k+\delta}\right]$$

To derive the aggregate portion of the equation, we must also derive asset holdings by groups: A_k.

To that end, recall that we have solved a differential equation for asset holdings at age τ:

$$a(\tau) = z_0\big[\psi_k(0)e^{\tilde{\mu}\tau} - \psi_k(\tau)\big]$$

Aggregate assets of type k individuals are given by integrating the over the age distribution and the skills of individuals:

$$\begin{aligned}
A_k &= \mathbb{E}[z_k|k]\int_0^\infty \big[\psi_k(0)e^{\tilde{\mu}\tau} - \psi_k(\tau)\big]\delta e^{-\delta\tau}d\tau \\
&= \delta\mathbb{E}[z_k|k]\left[\int_0^\infty \big[\psi_k(0)e^{\tilde{\mu}\tau}e^{-\delta\tau}d\tau - \int_0^\infty \psi_k(\tau)e^{-\delta\tau}d\tau\right] \\
&= \delta\mathbb{E}[z_k|k]\psi_k(0)\int_0^\infty e^{-(\delta-\tilde{\mu})\tau}d\tau - \delta\mathbb{E}[z_k|k]\int_0^\infty \psi_k(\tau)e^{-\delta\tau}d\tau \\
&= \delta\frac{\mathbb{E}[w_k\Psi_k z_k|k]}{r+\delta-g}\left[\frac{1}{\delta-\tilde{\mu}}\right] - \delta\mathbb{E}[z_k|k]\int_0^\infty \psi_k(\tau)e^{-\delta\tau}d\tau \\
&= \delta\frac{\mathbb{E}[w_k\Psi_k z_k|k]}{r+\delta-g}\left[\frac{1}{\delta-\tilde{\mu}}\right] - \frac{\mathbb{E}[w_k\Psi_k z_k|k]}{(r+\delta-g+\lambda_k\kappa_k)}\left[\frac{\kappa_k\lambda_k}{r+\delta-g} + \frac{\kappa_k\lambda_k+\delta}{\lambda_k+\delta}\right]
\end{aligned}$$

Noting that:

$$\begin{aligned}
\tilde{\mu} &= (r+\delta-g-\Gamma^{-\frac{1}{\sigma}}) \\
&= \frac{r-\rho-\sigma g}{\sigma}
\end{aligned}$$

,

We can then substitute this expression into the expression for C_k

$$C_k \equiv \Gamma^{-\frac{1}{\sigma}} \int_0^\infty \left(\mathbb{E}[a|k,\tau] + \psi_k(\tau)\mathbb{E}[z_k|k]\right) \delta \exp\{-\delta\tau\} d\tau$$

Given this, the expression for C_k is "easily" calculated as:

$$C_k = \Gamma^{-\frac{1}{\sigma}} \mathbb{E}[z_k|k] \delta \psi_k(0) \int_0^\infty e^{-(\delta-\mu)\tau} d\tau$$

Assuming that $\delta > \frac{r-\rho-\sigma g}{\sigma}$ this expression can be computed as:

$$\begin{aligned}
C_k &= \Gamma^{-\frac{1}{\sigma}} \mathbb{E}[z_k|k] \delta \psi_k(0) \left[\frac{1}{\delta - \frac{r-\rho-\sigma g}{\sigma}}\right] \\
&= \Gamma^{-\frac{1}{\sigma}} \mathbb{E}[z_k|k] \delta \psi_k(0) \left[\frac{\sigma}{\delta\sigma - r + \rho + \sigma g}\right] \\
&= \Gamma^{-\frac{1}{\sigma}} \mathbb{E}[w_k \Psi z_k|k] \frac{\delta}{r+\delta-g} \left[\frac{\sigma}{\delta\sigma - r + \rho + \sigma g}\right] \\
&= \Gamma^{-\frac{1}{\sigma}} \mathbb{E}[w_k \Psi z_k|k] \frac{\delta}{r+\delta-g} \left[\frac{\sigma}{\delta\sigma - r + \rho + \sigma g}\right] \\
&= \left[\frac{\rho + \delta + (\sigma-1)(r+\delta)}{\sigma}\right] \mathbb{E}[w_k \Psi z_k|k] \frac{\delta}{r+\delta-g} \left[\frac{\sigma}{\delta\sigma - r + \rho + \sigma g}\right] \\
&= \mathbb{E}[w_k \Psi z_k|k] \frac{\delta}{r+\delta-g} \left[\frac{\rho + \delta + (\sigma-1)(r+\delta)}{\delta\sigma - r + \rho + \sigma g}\right]
\end{aligned}$$

Where the third line follows from the fact that $\psi_k(0) = w_k \Psi_k \frac{1}{r+\delta-g}$

The last object required to characterize the Euler equation is:

$$\begin{aligned}
C_k(0) &= \Gamma^{-\frac{1}{\sigma}} \mathbb{E}[z_0|k] \psi_k(0) \\
C_k(0) &= \frac{\Gamma^{-\frac{1}{\sigma}}}{r+\delta-g} \mathbb{E}[w_k \Psi_k z_k|k]
\end{aligned}$$

This is convenient because recall that:

$$\forall k, k' \; \mathbb{E}[w_k \Psi_k z_k | k] = \mathbb{E}[w_{k'} \Psi_{k'} z_{k'} | k']$$

Note also that

$$\begin{aligned} \frac{C_k(0)}{C_k} &= \frac{\Gamma^{-\frac{1}{\sigma}} \mathbb{E}[z_0|k] \psi_k(0)}{\Gamma^{-\frac{1}{\sigma}} \mathbb{E}[z_k|k] \delta \psi_k(0) \left[\frac{\sigma}{\delta\sigma - r + \rho + \sigma g}\right]} \\ &= \frac{\delta\sigma - r + \rho + \sigma g}{\sigma\delta} \end{aligned}$$

Noting that

$$\delta \left[1 - \frac{C_k(0)}{C_k}\right] = \frac{r - \rho - \sigma g}{\sigma}$$

$$\begin{aligned} \frac{\dot{C}}{C} &= \left(\frac{r - \rho - \sigma g}{\sigma}\right) - \delta \sum_{k \in K} \frac{\mathbb{P}_k C_k}{C} \left[\frac{C_k - C_k(0)}{C_k}\right] \\ &= \left(\frac{r - \rho - \sigma g}{\sigma}\right) - \left(\frac{r - \rho - \sigma g}{\sigma}\right) \\ &= 0 \end{aligned}$$

Note that along the BGP the growth rate of normalized consumption is 0, and by the prior result C_k is the same across types.

C.2 The Log Utility Case

Here we present the case for log utility, which we use in the main model. For brevity, I solve the value function given an arbitrary future sequence of prices. This nests the balanced growth path

case where aggregates and real wages grow at a constant rate. Suppressing the functional form of $h_k(\tau)$, we can write the Hamilton-Jacobi-Bellman equation as:

$$(\rho+\delta)V_t(a,z,k,\tau) - \dot{V}_t = \max_c \left\{ \ln(c) + \frac{\partial V_t}{\partial a}\big[(r_t+\delta)a + w_{kt} z h_k(\tau) + \frac{\partial V_t}{\partial \tau} \right\} \tag{47}$$

The first step is to derive the normalized problem. To that end, let us define $\hat{a} = \frac{a}{Z_t}$, $\hat{c} = \frac{c}{Z_t}$ and $\hat{w}_{kt} = \frac{w_{kt}}{Z_t}$. Guess that the value function takes the form:

$$V_t = \frac{\ln(Z_t)}{\rho+\delta} + v_t(\tilde{a}, z, k, \tau)$$

Note the following:

$$\begin{aligned}
\frac{\partial V}{\partial a} &= \frac{\partial v}{\partial \hat{a}}\frac{\partial \hat{a}}{\partial a} = \frac{1}{Z_t}\frac{\partial v}{\partial \hat{a}} \\
\frac{\partial V}{\partial \tau} &= \frac{\partial v}{\partial \tau} \\
\frac{\partial V}{\partial t} &= \frac{g_t}{\rho+\delta} + \frac{\partial v}{\partial \hat{a}}\frac{\partial \hat{a}}{\partial t}\bigg|_a + \frac{\partial v}{\partial t} \\
&= \frac{g_t}{\rho+\delta} - \frac{\partial v}{\partial \hat{a}} g_t \hat{a} + \frac{\partial v}{\partial t}
\end{aligned}$$

Substituting into the HJB, we have that:

$$\begin{aligned}
\ln(Z_t) + (\rho+\delta)v_t(\tilde{a},z,k,\tau) - \frac{g_t}{\rho+\delta} - \frac{\partial v_t}{\partial t} + \frac{\partial v}{\partial \hat{a}} g_t \hat{a} = \\
\max_{\hat{c}} \left\{ \ln(Z_t) + \ln(\hat{c}) + \frac{1}{Z_t}\frac{\partial v}{\partial \hat{a}}\big[(r_t+\delta)Z_t\hat{a} - \hat{w}_{kt}Z_t z h_k(\tau) - Z_t\hat{c}\big] + \frac{\partial v_t}{\partial \tau} \right\}
\end{aligned}$$

Canceling $\ln(Z_t)$ from both sides and moving over the expression for $\frac{\partial v}{\partial \hat{a}} g_t \hat{a}$ delivers an expression for the normalized HJB.

$$(\rho+\delta)v_t(\hat{a},z,k,\tau)-\frac{g_t}{\rho+\delta}-\dot{v}_t=\max_{\hat{c}}\left\{\ln(\hat{c})+\frac{\partial v_t}{\partial \hat{a}}\left[(r_t+\delta-g_t)\hat{a}+\hat{w}_{kt}zh_k(\tau)-\hat{c}\right]+\frac{\partial v_t}{\partial \tau}\right\}$$

This is the normalized value for a type k worker. The next step is to solve for the closed form expression for this. To that end, guess that:

$$v_t(\hat{a},z,k,\tau)=B_t+\Gamma\ln(\hat{a}+\psi_k(t,\tau)z)$$

Note the following:

$$\begin{aligned}\frac{\partial v_t}{\partial t}&=\dot{B}_t+\frac{\Gamma}{\hat{a}+\psi_k(t,\tau)z}\frac{\partial \psi_k}{\partial t}\\ \frac{\partial v_t}{\partial \hat{a}}&=\frac{\Gamma}{\hat{a}+\psi_k(t,\tau)z}\\ \frac{\partial v_t}{\partial \tau}&=\frac{\Gamma}{\hat{a}+\psi_k(t,\tau)z}\frac{\partial \psi_k}{\partial \tau}\end{aligned}$$

Substituting into the HJB and using the optimality condition that $\frac{1}{\hat{c}}=\frac{\partial v_t}{\partial \hat{a}}$, we have that:

$$\begin{aligned}(\rho+\delta)\big(B_t+\Gamma\ln(\hat{a}+\psi_k(t,\tau)z)\big)-\frac{g_t}{\rho+\delta}-\dot{B}_t-\frac{\Gamma}{\hat{a}+\psi_k(t,\tau)z}\frac{\partial \psi_k}{\partial t}=\\ -\ln(\Gamma)+\ln(\hat{a}+\psi_k(t,\tau)z)+\frac{\Gamma}{\hat{a}+\psi_k(t,\tau)z}\left[(r_t+\delta-g_t)\hat{a}+\hat{w}_{kt}zh_k(\tau)-\hat{c}\right]+\frac{\Gamma}{\hat{a}+\psi_k(t,\tau)z}\frac{\partial \psi_k}{\partial \tau}\end{aligned}$$

Matching coefficients on $\ln(\hat{a}+\psi_k(t,\tau)z)$ delivers that:

$$\Gamma=\frac{1}{\rho+\delta}$$

Rewriting the prior terms after canceling the expression for $ln(\hat{a}+\psi_k(t,\tau)z)$ delivers that

$$(\rho+\delta)B_t - \frac{g_t}{\rho+\delta} - \dot{B}_t - \frac{1}{(\hat{a}+\psi_k(t,\tau)z)(\rho+\delta)} z\frac{\partial \psi_k}{\partial t} =$$
$$\ln(\rho+\delta) + \frac{1}{(\rho+\delta)(\hat{a}+\psi_k(t,\tau)z)}\big[(r_t+\delta-g_t)\hat{a} + \hat{w}_{kt}zh_k(\tau) - (\rho+\delta)(\hat{a}+\psi_k(t,\tau)z)\big] +$$
$$\frac{1}{(\rho+\delta)(\hat{a}+\psi_k(t,\tau)z)} z\frac{\partial \psi_k}{\partial \tau}$$

$$(\rho+\delta)B_t - \frac{g_t}{\rho+\delta} - \dot{B}_t - \ln(\rho+\delta) + 1 =$$
$$\frac{1}{(\rho+\delta)(\hat{a}+\psi_k(t,\tau)z)}\left[z\frac{\partial \psi_k}{\partial t} + z\frac{\partial \psi_k}{\partial \tau} + (r_t+\delta-g_t)\hat{a} + \hat{w}_{kt}zh_k(\tau)\right]$$

In order for this hold for $\hat{a}$, the expression on the interior cannot depend on a. This implies that:

$$\left[\frac{\partial \psi_k}{\partial t}z + \frac{\partial \psi_k}{\partial \tau}z + (r_t+\delta-g_t)\hat{a} + \hat{w}_{kt}zh_k(\tau)\right] = (r_t+\delta-g_t)(\hat{a}+\psi_k(t,\tau)z)$$

This delivers us a PDE for ψ_k:

$$\frac{\partial \psi_k}{\partial t}z + \frac{\partial \psi_k}{\partial \tau}z = (r_t+\delta-g_t)\psi_k(t,\tau)z - \hat{w}_{kt}zh_k(\tau)$$

As τ and t are collinear, this method can naturally be solved using the method of characteristics. Let s denote the change in calendar time. Then we can write:

$$\frac{d\psi_k}{ds} = \frac{\partial \psi_k}{\partial t}\frac{\partial t}{\partial s} + \frac{\partial \psi_k}{\partial \tau}\frac{\partial \tau}{\partial s} = \frac{\partial \psi_k}{\partial t} + \frac{\partial \psi_k}{\partial \tau}$$

This implies that we can write the PDE as an ODE in s, as:

$$\frac{d\psi_k}{ds} = (r_{t_0+s} + \delta - g_{t_0+s})\psi_k(t+s, \tau+s) - \hat{w}_{k,t_0+s}h_k(\tau+s)$$

This is linear differential equation with time varying coefficients. It can be solved via the method of integrating factors to give:

$$\psi_k(t,\tau) = \int_0^\infty \exp\left\{ -\int_0^s r_{t+u} + \delta - g_{t+u}du \right\} \hat{w}_{k,t+s}h_k(\tau+s)ds$$

Substituting in for the functional form of $h_k(\tau)$ gives the following expression for the value of a unit of human capital:

$$\begin{aligned}
\psi_k(t,\tau) &= \int_0^\infty \exp\left\{ -\int_0^s r_{t+u} + \delta - g_{t+u}du \right\} \hat{w}_{k,t+s}h_k(\tau+s)ds \\
&= \int_0^\infty \exp\left\{ -\int_0^s r_{t+u} + \delta - g_{t+u}du \right\} \hat{w}_{k,t+s}\left[\kappa_k - (\kappa_k - 1)\exp\left\{ -\lambda_k(\tau+s)\right\}\right]ds \\
&= \kappa_k \underbrace{\int_0^\infty \exp\left\{ -\int_0^s r_{t+u} + \delta - g_{t+u}du \right\} \hat{w}_{k,t+s}ds}_{\equiv \tilde{\mathbf{H}}_k(t)} \\
&- (\kappa_k - 1)e^{-\lambda_k\tau} \underbrace{\int_0^\infty \exp\left\{ -\int_0^s r_{t+u} + \delta - g_{t+u} + \lambda_k\, du \right\} \hat{w}_{k,t+s}ds}_{\hat{\mathbf{H}}_k(t)} \\
&= \kappa_k\tilde{\mathbf{H}}_k(t) - (\kappa_k - 1)e^{-\lambda_k\tau}\hat{\mathbf{H}}_k(t)
\end{aligned}$$

When we come to aggregation, solving the aggregate system will depend on solving for these variables. To that end, note that they satisfy the following differential equations:

$$\frac{d\tilde{\mathbf{H}}_k(t)}{dt} = (r_t + \delta - g_t)\tilde{\mathbf{H}}_k(t) - \hat{w}_k(t) \tag{48}$$

$$\frac{d\hat{\mathbf{H}}_k(t)}{dt} = (r_t + \delta - g_t + \lambda_k)\hat{\mathbf{H}}_k(t) - \hat{w}_k(t) \tag{49}$$

This characterizes $\psi_k(t,\tau)$. The last step in solving for the analytical expression is to derive the expression for B_t. To that end, observe that:

$$\begin{aligned}
(\rho+\delta)B_t - \frac{g_t}{\rho+\delta} - \dot{B}_t + 1 - \ln(\rho+\delta) &= \frac{r_t+\delta-g_t}{\rho+\delta} \\
(\rho+\delta)B_t - \dot{B}_t &= \ln(\rho+\delta) - 1 + \frac{g_t}{\rho+\delta} + \frac{r_t+\delta-g_t}{\rho+\delta} \\
(\rho+\delta)B_t - \dot{B}_t &= \ln(\rho+\delta) - 1 + \frac{r_t+\delta}{\rho+\delta} \\
\dot{B}_t - (\rho+\delta)B_t &= 1 - \ln(\rho+\delta) - \frac{r_t+\delta}{\rho+\delta}
\end{aligned}$$

This is a standard ODE that can be solved using the method of integrating factors. To that end:

$$\begin{aligned}
e^{-(\rho+\delta)t}\dot{B}_t - (\rho+\delta)e^{-(\rho+\delta)t}B_t &= e^{-(\rho+\delta)t}\left[1 - \ln(\rho+\delta) - \frac{r_t+\delta}{\rho+\delta}\right] \\
\frac{d}{dt}\left[e^{-(\rho+\delta)t}B_t\right] &= e^{-(\rho+\delta)t}\left[1 - \ln(\rho+\delta) - \frac{r_t+\delta}{\rho+\delta}\right] \\
\int_t^T \frac{d}{ds}\left[e^{-(\rho+\delta)s}B_s\right]ds &= \int_t^T e^{-(\rho+\delta)s}\left[1 - \ln(\rho+\delta) - \frac{r_{t+s}+\delta}{\rho+\delta}\right]ds \\
e^{-(\rho+\delta)T}B_t - e^{-(\rho+\delta)t}B_t &= \int_t^T e^{-(\rho+\delta)s}\left[1 - \ln(\rho+\delta) - \frac{r_{t+s}+\delta}{\rho+\delta}\right]ds
\end{aligned}$$

Imposing the condition that B_t is bounded from above taking the limit as $T \to \infty$ delivers that:

$$
\begin{aligned}
-e^{-(\rho+\delta)t}B_t &= \int_t^\infty e^{-(\rho+\delta)s}\left[1-\ln(\rho+\delta)-\frac{r_{t+s}+\delta}{\rho+\delta}\right]ds \\
&= \int_t^\infty e^{-(\rho+\delta)s}ds - \ln(\rho+\delta)\int_t^\infty e^{-(\rho+\delta)s}ds - \frac{1}{\rho+\delta}\int_t^\infty e^{-(\rho+\delta)s}(r_{t+s}+\delta)ds \\
B_t &= -\int_t^\infty e^{-(\rho+\delta)(s-t)}ds + \ln(\rho+\delta)\int_t^\infty e^{-(\rho+\delta)(s-t)}ds + \frac{1}{\rho+\delta}\int_t^\infty e^{-(\rho+\delta)(s-t)}(r_{t+s}+\delta)ds \\
B_t &= -\int_0^\infty e^{-(\rho+\delta)s}ds + \ln(\rho+\delta)\int_0^\infty e^{-(\rho+\delta)s}ds + \frac{1}{\rho+\delta}\int_0^\infty e^{-(\rho+\delta)(s)}(r_{t+s}+\delta)ds \\
B_t &= -\frac{1}{\rho+\delta} + \frac{\ln(\rho+\delta)}{\rho+\delta} + \frac{1}{\rho+\delta}\int_0^\infty e^{-(\rho+\delta)(s)}(r_{t+s}+\delta)ds \\
B_t &= \frac{\ln(\rho+\delta)-1}{\rho+\delta} + \frac{1}{\rho+\delta}\int_0^\infty e^{-(\rho+\delta)s}(r_{t+s}+\delta)ds
\end{aligned}
$$

Note that B_t does not depend on any k-specific items. This means that we can write[14] the choice-problem of a new-born as as:

$$N_t(z_0,\ldots,z_K) = \arg\max_k\{\psi_k(0,t)z_k\}$$

Before characterizing the choice problem, it is convenient to derive expressions for the aggregate distribution of efficiency units and for aggregate human wealth of type k. The expression for aggregate efficiency units of type k is given:

$$
\begin{aligned}
\mathbf{L}_{kt} &= \int_0^\infty \delta e^{-\delta\tau}\underbrace{\pi_k(\tau)\mathbb{E}[z_k|\text{ choose } k \text{ and age } \tau]}_{\equiv\eta_{kt}(\tau)}\left[\kappa_k-(\kappa_k-1)e^{-\lambda_k\tau}\right]d\tau \\
&= \kappa_k\underbrace{\int_0^\infty \delta e^{-\delta\tau}\eta_{kt}(\tau)d\tau}_{\equiv\tilde{\mathbf{L}}_{kt}} - (\kappa_k-1)\underbrace{\int_0^\infty \delta\eta_{kt}(\tau)e^{-(\delta+\lambda_k)\tau}d\tau}_{\equiv\hat{\mathbf{L}}_{kt}} \\
&= \kappa_k\tilde{\mathbf{L}}_{kt} - (\kappa_k-1)\hat{\mathbf{L}}_{kt}
\end{aligned}
$$

[14] After removing portions of the value that do not depend on the choice k.

Note that, $\tilde{\mathbf{L}}_{kt}$ and $\hat{\mathbf{L}}_{kt}$ both satisfy their own differential equations:

$$\frac{d\tilde{\mathbf{L}}_{\mathbf{kt}}}{dt} = \delta\eta_{kt}(0) - \delta\tilde{\mathbf{L}}_{\mathbf{kt}}$$
$$\frac{d\hat{\mathbf{L}}_{\mathbf{kt}}}{dt} = \delta\eta_{kt}(0) - (\delta + \lambda_k)\hat{\mathbf{L}}_{\mathbf{kt}}$$

Let $\mathbf{H}_{kt}$ denote the value of aggregate human wealth of type k. At any given point, this expression is given by:

$$\begin{aligned}
\mathbf{H}_{kt} &= \int_0^\infty \delta e^{-\delta\tau} \underbrace{\pi_k(\tau)\mathbb{E}[z_k|\text{ choose } k \text{ and age } \tau]}_{\equiv\eta_{kt}(\tau)} \psi_k(t,\tau)d\tau \\
&= \int_0^\infty \delta e^{-\delta\tau}\eta_{kt}(\tau)\psi_k(t,\tau)d\tau \\
&= \int_0^\infty \delta e^{-\delta\tau}\eta_{kt}(\tau)\left[\kappa_k\tilde{\mathbf{H}}_{kt} - (\kappa_k - 1)e^{-\lambda_k\tau}\hat{\mathbf{H}}_{kt}\right]d\tau \\
&= \kappa_k\tilde{\mathbf{H}}_{kt}\underbrace{\int_0^\infty \delta e^{-\delta\tau}\eta_{kt}(\tau)d\tau}_{\equiv\tilde{\mathbf{L}}_{kt}} - (\kappa_k - 1)\hat{\mathbf{H}}_{kt}\underbrace{\int_0^\infty \delta e^{-(\delta+\lambda_k)\tau}\eta_{kt}(\tau)d\tau}_{\equiv\hat{\mathbf{L}}_{kt}} \\
&= \kappa_k\tilde{\mathbf{H}}_{kt}\tilde{\mathbf{L}}_{\mathbf{kt}} - (\kappa_k - 1)\hat{\mathbf{H}}_{kt}\hat{\mathbf{L}}_{kt}
\end{aligned}$$

Note also that the linearity of the consumption policy function implies that we can write aggregate normalized consumption as:

$$\mathbf{C}_t = (\rho + \delta)(\mathbf{A}_t + \mathbf{H}_t)$$

The last step in characterizing the household side of the model is to characterize the selection process for new entrants. This requires specifying a distribution of initial productivity draws. As in the main-text, I assume that the distribution of productivity draws is multivariate frechet with a cross-nested CES correlation function. That is to say that the correlation function is given by:

$$G(x_0, \ldots, x_K) = \sum_{f \in F} \left(\sum_{k \in K} (\omega_{fk} x_k)^{\frac{1}{1-\rho_f}} \right)^{1-\rho_f}$$

Recall that the choice problem of a new-born is equivalent to the following:

$$N_t(z_0, \ldots, z_K) = \arg\max_k \{\psi_k(0,t) z_k\}$$

And that $\psi_k(0,t) = \kappa_k \tilde{\mathbf{H}}_k(t) - (\kappa_k - 1)\hat{\mathbf{H}}_k(t)$. The consequence of this is that we do not need to separately track the human wealth of a new entrant choosing type k. Rather this quantity can be derived analytically given the state variables $\tilde{\mathbf{H}}_k(t)$ and $\hat{\mathbf{H}}_k(t)$. Appendix D provides a number of results on multivariate-Frechet distributions. Those results allow us to characterize in closed form the flow of efficiency units into each field.

C.3 Training Periods

The baseline quantitative model includes an additional period of training before human capital begins to accrue. This is designed to replicate the PhD training period. In this training period, I assume that agents are hand-to-mouth and do not accumualate experience. Individuals transition to fully-fledged graduates according to an exponential random variable μ. Let $N_t(z,k)$ denote the time t value of a trainee having chosen k who has innate skill z. Given this, we can write the the value of a worker upon receipt of the PhD, we can write the HJB of a person having chosen to become a trainee of type k as:

$$(\rho+\delta)N_t(z,k) - \dot{N}_t = \frac{(w_{kt}z_k)^{1-\sigma}}{1-\sigma} + \mu_k\big(V_t(0,z,k,0) - N_t(z,k)\big)$$

C.3.1 Along the BGP

Along the BGP, aggregates grow at a constant rate g. To characterize N, we will proceed by guessing that:

- $\dot{N}_t = (1-\sigma)gN_t$
- $N_t = z^{1-\sigma}n_t$

Using the expression for $V(0,z,0,k)$ derived in the prior section, we can thus write the whole HJB as:

$$(\rho+\delta+\mu_k-(1-\sigma)g)N_t(z,k) = \frac{(w_k z)^{1-\sigma}}{1-\sigma} + \frac{\mu_k}{1-\sigma}\left[\frac{(\rho+\delta)+(\sigma-1)(r+\delta)}{\sigma}\right]^{-\sigma}\left(w_k z\left[\frac{r+\delta-g+\kappa_k\lambda_k}{r+\delta-g+\lambda_k}\right]\right)^{1-\sigma}$$

$$(\rho+\delta+\mu_k-(1-\sigma)g)z^{1-\sigma}n_t = \frac{(w_{kt}z)^{1-\sigma}}{1-\sigma}+ \frac{\mu_k}{1-\sigma}\left[\frac{(\rho+\delta)+(\sigma-1)(r+\delta)}{\sigma}\right]^{-\sigma}\left(w_k z\left[\frac{r+\delta-g+\kappa_k\lambda_k}{r+\delta-g+\lambda_k}\right]\right)^{1-\sigma}$$

$$\begin{aligned}(\rho+\delta+\mu_k-(1-\sigma)g)z^{1-\sigma}n_t &= \frac{(w_{kt}z)^{1-\sigma}}{1-\sigma}+ \frac{\mu_k}{1-\sigma}\left[\frac{(\rho+\delta)+(\sigma-1)(r+\delta)}{\sigma}\right]^{-\sigma}\left(w_k z\left[\frac{r+\delta-g+\kappa_k\lambda_k}{r+\delta-g+\lambda_k}\right]\right)^{1-\sigma}\\ (\rho+\delta+\mu_k-(1-\sigma)g)n_t &= \frac{(w_{kt})^{1-\sigma}}{1-\sigma}+ \frac{\mu_k}{1-\sigma}\left[\frac{(\rho+\delta)+(\sigma-1)(r+\delta)}{\sigma}\right]^{-\sigma}\left(w_k \left[\frac{r+\delta-g+\kappa_k\lambda_k}{r+\delta-g+\lambda_k}\right]\right)^{1-\sigma}\end{aligned}$$

Doing further algebra implies that:

$$n = \frac{w_k^{1-\sigma}}{(1-\sigma)\big(\rho+\delta+(1-\sigma)g)\big)}\left[1+\mu_k\left[\frac{(\rho+\delta)+(\sigma-1)(r+\delta)}{\sigma}\right]^{-\sigma}\left(\left[\frac{r+\delta-g+\kappa_k\lambda_k}{r+\delta-g+\lambda_k}\right]\right)^{1-\sigma}\right]$$

C.3.2 The Log Case

Consider now the limit case as $\sigma \to 1$. This case greatly simplifies the computation of the equilibrium. The HJB in this case is given by:

$$(\rho+\delta)N_t(z,k) - \dot{N}_t = \ln(w_{kt}z_k) + \mu_k\big(V_t(0,z,k,0) - N_t(z,k)\big)$$

As in the prior step, normalize we proceed first by normalizing the value function. In this case, we guess that the value function has form: $N_t(z,k) = \frac{\ln(Z_t)}{\rho+\delta} + \hat{N}_t(z,k)$. This implies that:

$$\frac{\partial N_t}{\partial t} = \frac{g_t}{\rho + \delta} + \dot{\tilde{N}}_t$$

Substituting into the value function gives:

$$\begin{aligned}
\frac{\mu_k}{\rho+\delta}\ln(Z_t) + \ln(Z_t) + (\rho + \delta + \mu_k) n_t(z,k) - \frac{g_t}{\rho+\delta} - \dot{n}_t \\
= \ln(Z_t) + \ln(\hat{w}_{kt} z) + \mu_k \frac{\ln(Z_t)}{\rho+\delta} + \mu_k v_t(0,z,0,k) \\
(\rho + \delta + \mu_k) n_t(z,k) - \frac{g_t}{\rho+\delta} - \dot{n}_t = \ln(\hat{w}_{kt} z) + \mu_k v_t(0,z,0,k)
\end{aligned}$$

This is the normalized HJB of a new-born immediately upon choosing type k. It turns out that this problem can also be solved in closed form. To do so, guess that the value function takes the form:

$$n_t(z,k) = \Lambda_{kt} + \Theta \ln(z)$$

This implies that:

$$\dot{n}_t = \dot{\Lambda}_{kt}$$

Substituting in for n_t delivers:

$$(\rho + \delta + \mu_k)\big(\Lambda_{kt} + \Theta \ln(z)\big) - \frac{g_t}{\rho+\delta} - \dot{\Lambda}_{kt} = \ln(\hat{w}_{kt}) + \ln(z) + \mu_k (B_t + \frac{\ln(\psi_k(0,t))}{\rho+\delta} + \frac{\ln(z)}{\rho+\delta})$$

Matching coefficients on $\ln(z)$ delivers that:

$$\begin{aligned}(\rho+\delta+\mu_k)\Theta &= 1+\frac{\mu_k}{\rho+\delta}\\ &= \frac{\rho+\delta+\mu_k}{\rho+\delta}\\ \Theta &= \frac{1}{\rho+\delta}\end{aligned}$$

This implies that Λ_t is the function that solves the differential equation given by:

$$(\rho+\delta+\mu_k)\Lambda_{kt}-\dot{\Lambda}_{kt}=\frac{g_t}{\rho+\delta}+\ln(\hat{w}_{kt})+\mu_k B_t+\frac{\mu_k}{\rho+\delta}\ln(\psi_k(0,t))$$

To characterize the value of a newborn here requires one to solve 2 additiona ODEs, one for B_t and one for Λ_{kt}. Note, however, that the equation for B_t denotes the present value of growth, which is common across types. As a consequence, one needs to solve only an additional $K+1$ additional ODEs to characterize the value of a newborn. Conditional on these values, one can write the value of a newborn as:

$$N_{kt}=\arg\max_k\left\{\Lambda_{kt}+\frac{1}{\rho+\delta}\ln(z)\right\}$$

Given my assumption that the distribution of z is distributed multivariate Frechet, these choice probabilities in closed form. To see this recall that if z is distributed Frechet, then $\ln(z)$ is distributed Gumbel. As shown in Appendix D, the analogous result carries through in the multivariate Frechet case. The consequence of this is that one retains closed form expressions for the value of the maximum and of the choice probabilities. This allows me to write another differential equation for the quantity of training skill.

The next step is to derive the market clearing conditions in the economy with training periods. I firstly derive the condition for aggregate labor supply of trained researchers. This quantity is given by:

$$\begin{aligned}\mathbf{L}_{kt} &= \int_0^\infty \delta\mu_k \mathbf{E}_{k,t-\tau}\left[\kappa_k - (\kappa_k - 1)e^{-\lambda_k \tau}\right] e^{-\delta\tau} d\tau \\ &= \kappa_k \int_0^\infty \delta\mu_k \mathbf{E}_{k,t-\tau} e^{-\delta\tau} - (\kappa_k - 1)\int_0^\infty \delta\mu_k \mathbf{E}_{k,t-\tau} e^{-(\delta+\lambda_k)\tau} \\ &= \kappa_k \tilde{\mathbf{L}}_{kt} - (\kappa_k - 1)\hat{\mathbf{L}}_{kt}\end{aligned}$$

Analogously, for the human wealth of individuals, we have:

$$\mathbf{H}_{kt} = \kappa_k \tilde{\mathbf{H}}_{kt}\tilde{\mathbf{L}}_{kt} - (\kappa_k - 1)\hat{\mathbf{H}}_{kt}\hat{\mathbf{L}}_{kt}$$

Where $\tilde{\mathbf{H}}_{kt}$ and $\hat{\mathbf{H}}_{\mathbf{kt}}$ are defined analogously to before.

Aggregate labor supply in efficiency units is thus given by:

$$\mathcal{L}_{kt} = \mathbf{E}_{kt} + \mathbf{L}_{kt}$$

We track this by solving a set of ODEs. These are:

$$\begin{aligned}\frac{d\mathbf{E}_{kt}}{dt} &= \delta\eta_{kt} - (\delta + \mu_k)\mathbf{E}_{kt} \\ \frac{d\tilde{\mathbf{L}}_{kt}}{dt} &= \delta\mu_k \mathbf{E}_{kt} - \delta\hat{\mathbf{L}}_{kt} \\ \frac{d\hat{\mathbf{L}}_{kt}}{dt} &= \delta\mu_k \mathbf{E}_{kt} - (\delta + \lambda_k)\hat{\mathbf{L}}_{kt}\end{aligned}$$

Aggregate consumption in normalized terms becomes:

$$\mathbf{C}_t = \sum_k \hat{w}_{kt}\mathbf{E}_{kt} + (\rho + \delta)(\mathbf{A}_t + \mathbf{H}_t) \tag{50}$$

D Frechet Results

First I reproduce some useful results from before:

Lemma: Suppose that $z_i \sim Frechet(T_i, \kappa)$. *Then* $\nu_i z_i \sim Frechet(\nu_i^\kappa T_i, \kappa)$.

Proof. Suppose that $z_i \sim Frechet(T_i, \kappa)$. Then

$$\mathbb{P}[z_i \leq Z] = \exp\left\{ -(\frac{1}{T_i}z)^{-\kappa} \right\}$$

Note that $\mathbb{P}[\nu_i z_i \leq z] \iff \mathbb{P}[z_i \leq \frac{z}{\nu_i}]$

But:

$$\mathbb{P}[z_i \leq \frac{Z}{\nu_i}] = \exp\left\{ -(\frac{z}{T_i \nu_i})^{-\kappa} \right\}$$

Which is $Frechet(\nu_i^\kappa T_i, \kappa)$ ■

Lemma: Suppose that $z_i \sim Frechet(T_i, \kappa)\ \forall i$ *and that* G *is a max-stable correlation function. Then* $\ln(z_i) \sim Gumbel(\frac{\ln(T_i)}{\kappa}, \frac{1}{\kappa})$ *and the joint distribution also has correlation function* G*.*

Proof. The proof of the first part is standard so I neglect it here. The proof of the second part follows by inspection. Specifically note that:

$$F(z_1, z_2, \ldots, z_K) = \exp\left\{ - G(T_1 z_1^{-\kappa}, T_2 z_2^{-\kappa}, \ldots, T_K z_K^{-\kappa}) \right\}$$

Now apply a change of variables:

$$x_i = \ln(z_i) \iff z_i = e^{-x_i}$$

The joint distribution is then given by:

$$\begin{aligned}\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_K \leq x_K) &= \mathbb{P}(\ln(z_1) \leq x_1, \ldots, \ln(z_K) \leq x_K) \\ &= \exp\left\{ - G(T_1 e^{-\kappa x_1}, T_2 e^{-\kappa x_2}, \ldots, T_K e^{-\kappa x_K} \right\}\end{aligned}$$

Note that this is a multivariate Gumbel distribution with marginals that are gumbel and have location parameter $\frac{\ln(T_k)}{\kappa}$ and scale parameter $\frac{1}{\kappa}$. ■

Lemma: Suppose that $z_i \sim Frechet(T_i, \kappa)\ \forall i$ and that G is any correlation function. Then population proportions are given by:

$$\pi_k = \frac{T_k G_k(T_0, \ldots, T_K)}{G(T_0, \ldots, T_K)}$$

Proof.

$$\begin{aligned}
\pi_k &= \mathbb{P}\left[\max_{k'} z_{k'} \leq Z \ \wedge \max_{k'} z_{k'} = z_k\right] \\
&= \mathbb{P}\left[z_{k''} \leq Z \ \forall \ k'' \ \wedge z_{k'} \leq z_k \forall k'\right] \\
&= \mathbb{P}\left[z_{k'} \leq z_k \forall k' \wedge z_k \leq Z\right] \\
&= \int_0^Z \frac{\partial}{\partial z_k} \mathbb{P}[z_{k'} \leq z_k \forall k' \wedge z_k \leq t]\Big|_{\mathbf{z}=\iota t} dt \\
&= \int_0^Z \frac{\partial}{\partial z_k} \exp\left\{ -G(T_1 z_1^{-\kappa}, \ldots, T_K z_K^{-\kappa})\right\}\Big|_{\mathbf{z}=\iota t} dt \\
&= \int_0^Z \exp\left\{ -G(T_1 z_1^{-\kappa}, \ldots, T_K z_K^{-\kappa})\right\} G_k(T_1 z_1^{-\kappa}, \ldots, T_K z_K^{-\kappa}) \kappa T_k z_k^{-\kappa-1}\Big|_{\mathbf{z}=\iota t} dt \\
&= \int_0^Z \exp\left\{ -G(T_1 t^{-\kappa}, \ldots, T_K t^{-\kappa})\right\} G_k(T_1 t^{-\kappa}, \ldots, T_K t^{-\kappa}) \kappa T_k t^{-\kappa-1} dt \\
&= \int_0^Z \exp\left\{ -t^{-\kappa} G(T_1, \ldots, T_K)\right\} G_k(T_1, \ldots, T_K) \kappa T_k t^{-\kappa-1} dt \\
&= T_k G_k(T_1, \ldots, T_K) \int_0^Z \exp\left\{ -t^{-\kappa} G(T_1, \ldots, T_K)\right\} \kappa t^{-\kappa-1} dt \\
&= \frac{T_k G_k(T_1, \ldots, T_K)}{G(T_1, \ldots, T_K)} \int_0^Z \exp\left\{ -t^{-\kappa} G(T_1, \ldots, T_K)\right\} \kappa t^{-\kappa-1} G(T_1, \ldots, T_K) dt \\
&= \frac{T_k G_k(T_1, \ldots, T_K)}{G(T_1, \ldots, T_K)} \int_{-\infty}^{Z^{-\kappa} G(T_1, \ldots, T_K)} \exp\{-u\} du \\
&= \frac{T_k G_k(T_1, \ldots, T_K)}{G(T_1, \ldots, T_K)} \int_{-\infty}^{Z^{-\kappa} G(T_1, \ldots, T_K)} \exp\{-u\} du \\
&= \frac{T_k G_k(T_1, \ldots, T_K)}{G(T_1, \ldots, T_K)} \exp\left\{ -Z^{-\kappa} G(T_1, \ldots, T_K)\right\}
\end{aligned}$$

Taking limits as Z goes to infinity delivers the result. ■

Lemma: Suppose that $z_i \sim Frechet(T_i, \kappa)\ \forall i$ and that G is a correlation function. Then:

$$\mathbb{E}[Z_i | Z_i = \max_k Z_k] =$$

Proof.

$$\begin{aligned}
\mathbb{E}[Z_i|Z_i = \max_k Z_k] &= \int_0^\infty z\mathbb{P}[Z_i = z|Z_i = \max_k Z_k]dz \\
&= \int_0^\infty z\frac{\mathbb{P}[Z_i = \max_k Z_K|Z_i = z]}{\mathbb{P}[Z_i = \max_k Z_K]}f_i(z)dz \\
&= \frac{1}{\pi_i}\int_0^\infty z\mathbb{P}[Z_i = \max_k Z_K|Z_i = z]f_i(z)dz \\
&= \frac{1}{\pi_i}\int_{\Omega_z} \mathbb{P}[Z_i = \max_k Z_k|Z_i = z]zf_i(z)dz \\
&= \frac{1}{\pi_i}\int_0^\infty z\mathbb{P}[Z_k \leq Z_i \forall k \neq i\ \wedge Z_i = z]dz \\
&= \frac{1}{\pi_i}\int_0^\infty z\frac{\partial}{\partial Z_i}\mathbb{P}[Z_k \leq Z_i \forall k \neq i\ \wedge Z_i \leq z]dz \\
&= \frac{1}{\pi_i}\int_0^\infty z\frac{\partial}{\partial Z_i}\exp\Big\{-G(T_1Z_1^{-\theta}, \ldots, T_KZ_K^{-\theta})\Big\}\Big|_{\mathbf{Z}=\iota z}dz \\
&= \frac{1}{\pi_i}\int_0^\infty zG_i(T_1z^{-\theta}, \ldots, T_Kz^{-\theta})\exp\Big\{-G(T_1z^{-\theta}, \ldots, T_Kz^{-\theta})\Big\}\theta z^{-\theta-1}dz \\
&= \frac{1}{\pi_i}\int_0^\infty zT_iG_i(T_1, \ldots, T_K)\exp\Big\{-G(T_1, \ldots, T_K)z^{-\theta}\Big\}\theta z^{-\theta-1}dz \\
&= \frac{1}{\pi_i}T_iG_i(T_1, \ldots, T_K)\int_0^\infty z\exp\Big\{-G(T_1, \ldots, T_K)z^{-\theta}\Big\}\theta z^{-\theta-1}dz \\
&= \frac{1}{\pi_i}\underbrace{T_i\frac{G_i(T_1, \ldots, T_K)}{G(T_1, \ldots, T_K)}}_{\equiv\pi_i}\int_0^\infty zG(T_1, \ldots, T_K)\exp\Big\{-G(T_1, \ldots, T_K)z^{-\theta}\Big\}\theta z^{-\theta-1}dz
\end{aligned}$$

Now note that if $X \sim Frechet(G(T_1, \ldots, T_K), \theta)$, then

$$f_X(x) = G(T_1, \ldots, T_K)\exp\Big\{-G(T_1, \ldots, T_K)x^{-\theta}\Big\}\theta x^{-\theta-1}$$

As a consequence of this, we can write:

$$\begin{aligned}
\mathbb{E}[Z_i|Z_i = \max_k Z_k] &= \Gamma\left(1-\frac{1}{\theta}\right)G(T_1, \ldots, T_K)^{\frac{1}{\theta}} \\
&= \Gamma\left(1-\frac{1}{\theta}\right)\left(\frac{G(T_1, \ldots, T_K)}{T_iG_i(T_1, \ldots, T_K)}\right)^{\frac{1}{\theta}}(T_iG_i(T_1, \ldots, T_K))^{\frac{1}{\theta}} \\
&= \Gamma\left(1-\frac{1}{\theta}\right)\pi_i^{-\frac{1}{\theta}}(T_iG_i(T_1, \ldots, T_K))^{\frac{1}{\theta}}
\end{aligned}$$

■

Lemma: Suppose that $z_i \sim Frechet(T_i, \kappa)\ \forall i$ and that G is a correlation function. Then:

$$\mathbb{E}[Z_i^l | Z_i = \max_k Z_k] =$$

Proof.

$$\begin{aligned}
\mathbb{E}[Z_i^l | Z_i = \max_k Z_k] &= \int_0^\infty z^l \mathbb{P}[Z_i = z | Z_i = \max_k Z_k] dz \\
&= \int_0^\infty z^l \frac{\mathbb{P}[Z_i = \max_k Z_K | Z_i = z]}{\mathbb{P}[Z_i = \max_k Z_K]} f_i(z) dz \\
&= \frac{1}{\pi_i} \int_0^\infty z^l \mathbb{P}[Z_i = \max_k Z_K | Z_i = z] f_i(z) dz \\
&= \frac{1}{\pi_i} \int_{\Omega_z} \mathbb{P}[Z_i = \max_k Z_k | Z_i = z] z^l f_i(z) dz \\
&= \frac{1}{\pi_i} \int_0^\infty z^l \mathbb{P}[Z_k \leq Z_i \forall k \neq i \ \wedge Z_i = z] dz \\
&= \frac{1}{\pi_i} \int_0^\infty z^l \frac{\partial}{\partial Z_i} \mathbb{P}[Z_k \leq Z_i \forall k \neq i \ \wedge Z_i \leq z] dz \\
&= \frac{1}{\pi_i} \int_0^\infty z^l \frac{\partial}{\partial Z_i} \exp\Big\{ - G(T_1 Z_1^{-\theta}, \ldots, T_K Z_K^{-\theta}) \Big\} \Big|_{\mathbf{Z} = \iota z} dz \\
&= \frac{1}{\pi_i} \int_0^\infty z^l G_i(T_1 z^{-\theta}, \ldots, T_K z^{-\theta}) \exp\Big\{ - G(T_1 z^{-\theta}, \ldots, T_K z^{-\theta}) \Big\} \theta z^{-\theta-1} dz \\
&= \frac{1}{\pi_i} \int_0^\infty z^l T_i G_i(T_1, \ldots, T_K) \exp\Big\{ - G(T_1, \ldots, T_K) z^{-\theta} \Big\} \theta z^{-\theta-1} dz \\
&= \frac{1}{\pi_i} T_i G_i(T_1, \ldots, T_K) \int_0^\infty z^l \exp\Big\{ - G(T_1, \ldots, T_K) z^{-\theta} \Big\} \theta z^{-\theta-1} dz \\
&= \frac{1}{\pi_i} \underbrace{T_i \frac{G_i(T_1, \ldots, T_K)}{G(T_1, \ldots, T_K)}}_{\equiv \pi_i} \int_0^\infty z^l G(T_1, \ldots, T_K) \exp\Big\{ - G(T_1, \ldots, T_K) z^{-\theta} \Big\} \theta z^{-\theta-1} dz
\end{aligned}$$

Let $t \equiv G(T_1, \ldots, T_K) z^{-\theta}$. Then we have

$$dt = -G(T_1, \ldots, T_K) \theta z^{-\theta-1} \wedge z = \left[\frac{G(T_1, \ldots, T_K)}{t} \right]^{\frac{1}{\theta}}$$

. This implies that:

$$\begin{aligned}
\mathbb{E}[Z_i^l | Z_i = \max_k Z_k] &= \int_\infty^0 -G(T_1, \ldots, T_K)^{\frac{l}{\theta}} t^{-\frac{l}{\theta}} \exp(-t) dt \\
&= \int_0^\infty G(T_1, \ldots, T_K)^{\frac{l}{\theta}} t^{-\frac{l}{\theta}} \exp(-t) dt \\
&= G(T_1, \ldots, T_K)^{\frac{l}{\theta}} \int_0^\infty t^{-\frac{l}{\theta}} \exp(-t) dt \\
&= G(T_1, \ldots, T_K)^{\frac{l}{\theta}} \Gamma\left(1 - \frac{l}{\theta}\right) \textit{ so long as the integral converges} \\
&= T_i^{\frac{1}{\theta}} G_i(T_0, T_1, \ldots, T_K)^{\frac{l}{\theta}} \pi_i^{-\frac{l}{\theta}} \Gamma\left(1 - \frac{l}{\theta}\right)
\end{aligned}$$

■

***Lemma**: Let $Z_1, Z_2, \ldots, Z_K$ be multivariate Frechet with scale parameters T_i, shape parameter θ, and correlation function $G(x_1, \ldots, x_K)$. Then $\max_i\{\nu_i Z_i\}$ is Frechet with shape parameter α and scale parameter $G(\nu_1^\theta T_1, \ldots, \nu_K^\theta T_K)$.*

Proof.

$$\begin{aligned}
\mathbb{P}[\max_k \nu_k Z_k \leq t] &= \mathbb{P}[\nu_i z_i \leq t \ \forall i] \\
&= \mathbb{P}[Z_i \leq \frac{t}{\nu_i} \ \forall i] \\
&= \exp\left\{-G(T_1\nu_1^\theta t^{-\theta}, \ldots, T_K\nu_K^\theta t^{-\theta})\right\} \\
&= \exp\left\{-G(T_1\nu_1^\theta, \ldots, T_K\nu_K^\theta) t^{-\theta}\right\}
\end{aligned}$$

Which is the CDF of a Frechet with scale parameter $G(T_1\nu_1^\theta, \ldots, T_K\nu_K^\theta)$ and shape parameter θ. ■

***Lemma**: $Z_1, Z_2, \ldots, Z_K$ be multivariate Frechet with scale parameters T_i, shape parameter θ, and correlation function $G(x_1, \ldots, x_K)$. Then $\mathbb{E}[\nu_i z_i |$ choose $i] = \mathbb{E}[\max_k \nu_k z_k]$.*

Proof.

$$
\begin{aligned}
\mathbb{E}[\nu_i z_i | \text{ choose i}] &= \mathbb{E}[\nu_i z_i | \nu_i z_i \geq \nu_k z_k \ \forall k \neq i] \\
&= \int_0^\infty t \mathbb{P}[\nu_i z_i = t | \nu_i z_i \geq \nu_k z_k \ \forall k \neq i] dt \\
&= \frac{1}{\pi_i} \int_0^\infty t \mathbb{P}[\nu_i z_i = t \wedge \nu_i z_i \geq \nu_k z_k \ \forall k \neq i] dt \\
&= \frac{1}{\pi_i} \int_0^\infty t \frac{\partial}{\partial Z_i} \mathbb{P}[\nu_i z_i \leq t \wedge \nu_k z_k \leq \nu_i z_i] \Big|_{z_i = \frac{t}{\nu_i}} dt \\
&= \frac{1}{\pi_i} \int_0^\infty t \frac{\partial}{\partial Z_i} \mathbb{P}[\max_k \nu_k Z_k \leq t] \Big|_{\nu \mathbf{z} = \iota t} dt \\
&= \frac{1}{\pi_i} \int_0^\infty t \frac{\partial}{\partial Z_i} \exp\left\{-G(T_1 \nu_1^\theta t^{-\theta}, \ldots, T_K \nu_K^\theta t^{-\theta})\right\} dt \\
&= \frac{1}{\pi_i} \int_0^\infty t \frac{\partial}{\partial Z_i} \exp\left\{-G(T_1 \nu_1^\theta t^{-\theta}, \ldots, T_K \nu_K^\theta t^{-\theta})\right\} dt \\
&= \frac{1}{\pi_i} \int_0^\infty t\theta t^{-\theta-1} T_i \nu_i^\theta G_i(T_1 \nu_1 t^{-\theta}, \ldots, T_K \nu_K t^{-\theta}) \exp\left\{-G(T_1 \nu_1^\theta t^{-\theta}, \ldots, T_K \nu_K^\theta t^{-\theta})\right\} dt \\
&= \frac{1}{\pi_i} T_i \nu_i^\theta G_i(T_1 \nu_1 t^{-\theta}, \ldots, T_K \nu_K t^{-\theta}) \int_0^\infty t\theta t^{-\theta-1} \exp\left\{-G(T_1 \nu_1^\theta t^{-\theta}, \ldots, T_K \nu_K^\theta t^{-\theta})\right\} dt \\
&= \frac{1}{\pi_i} \frac{T_i \nu_i^\theta G_i(T_1 \nu_1 t^{-\theta}, \ldots, T_K \nu_K t^{-\theta})}{G(T_1 \nu_1^\theta, \ldots, T_K \nu_K^\theta)} \int_0^\infty t\theta t^{-\theta-1} G(T_1 \nu_1^\theta, \ldots, T_K \nu_K^\theta) \exp\left\{-G(T_1 \nu_1^\theta t^{-\theta}, \ldots, T_K \nu_K^\theta t^{-\theta})\right\} dt \\
&= \int_0^\infty t\theta t^{-\theta-1} G(T_1 \nu_1^\theta, \ldots, T_K \nu_K^\theta) \exp\left\{-G(T_1 \nu_1^\theta t^{-\theta}, \ldots, T_K \nu_K^\theta t^{-\theta})\right\} dt \\
&= \int_0^\infty t\theta t^{-\theta-1} G(T_1 \nu_1^\theta, \ldots, T_K \nu_K^\theta) \exp\left\{-G(T_1 \nu_1^\theta, \ldots, T_K \nu_K^\theta) t^{-\theta}\right\} dt \\
&= \int_\infty^0 -G(T_1 \nu_1^\theta, \ldots, T_K \nu_K^\theta)^{\frac{1}{\theta}} u^{-\frac{1}{\theta}} \exp\{-u\} du \\
&= G(T_1 \nu_1^\theta, \ldots, T_K \nu_K^\theta)^{\frac{1}{\theta}} \int_0^\infty u^{-\frac{1}{\theta}} \exp\{-u\} du \\
&= G(T_1 \nu_1^\theta, \ldots, T_K \nu_K^\theta)^{\frac{1}{\theta}} \Gamma\left(1 - \frac{1}{\theta}\right) = \mathbb{E}[\max_k \nu_k z_k] \text{ by the prior lemma.}
\end{aligned}
$$

■

E Estimation

This appendix contains supporting algebra, plots, and details about the estimation strategy used to find the parameters characterizing the labor supply system.

E.1 Derivation of Elasticities

Note that I neglect the arguments of ψ_k, and abuse notation by writing $\psi_k = \psi_k(t, 0)$.

$$\sum_{f \in F} \left[\sum_{k \in K} \left(\omega_{kf} T_k \psi_k^{\theta} \right)^{\frac{1}{1-\rho_f}} \right]^{1-\rho_f}$$

From here, we need to derive the population proportions and average earnings. Let us begin with population proportions.

Using lemmas derived in Appendix D, we begin by differentiating the correlation function with respect to argument i.

$$G_i(x_0, \dots x_k) = \sum_f \left\{ \left[\sum_k (\omega_{fk} x_k)^{\frac{1}{1-\rho_f}} \right]^{1-\rho_f} \times \frac{(\omega_{fk} x_k)^{\frac{1}{1-\rho_f}-1} \omega_{fk}}{\sum_k (\omega_{fk} x_k)^{\frac{1}{1-\rho_f}}} \right\}$$

It follows that one can write:

$$\begin{aligned}
\mathbb{P}_k &= \frac{x_k G_i(x_0, \dots, x_K)}{G(x_0, \dots, x_K)} \\
&= \sum_f \left\{ \underbrace{\frac{\left[\sum_k (\omega_{fk} x_k)^{\frac{1}{1-\rho_f}} \right]^{1-\rho_f}}{\sum_f \left[\sum_k (\omega_{fk} x_k)^{\frac{1}{1-\rho_f}} \right]^{1-\rho_f}}}_{\equiv \mathbb{P}_f} \times \underbrace{\frac{(\omega_{fk} x_k)^{\frac{1}{1-\rho_f}}}{\sum_k (\omega_{fk} x_k)^{\frac{1}{1-\rho_f}}}}_{\equiv \mathbb{P}_{k|f}} \right\} \\
\mathbb{P}_k &= \sum_f \mathbb{P}_{k|f} \mathbb{P}_f
\end{aligned}$$

Substituting in for the the definition of the marginal Frechet variables, we have:

$$\mathbb{P}_k = \sum_f \left\{ \underbrace{\frac{\left[\sum_{k'} (\omega_{fk'} T_{k'} \psi_{k'}^{\theta})^{\frac{1}{1-\rho_f}}\right]^{1-\rho_f}}{\sum_f \left[\sum_{k'} (\omega_{fk'} T_{k'} \psi_{k'}^{\theta})^{\frac{1}{1-\rho_f}}\right]^{1-\rho_f}}}_{\equiv \mathbb{P}^f} \times \underbrace{\frac{(\omega_{fk} T_k \psi_k^{\theta})^{\frac{1}{1-\rho_f}}}{\sum_{k'} (\omega_{fk'} T_{k'} \psi_{k'}^{\theta})^{\frac{1}{1-\rho_f}}}}_{\equiv \mathbb{P}_{k|f}} \right\}$$

It is convenient to write the derivative of the choice probabilities and perform a set of manipulations:

$$\begin{aligned}
\frac{\partial \mathbb{P}_k}{\partial \psi_i} &= \sum_f \frac{\partial \mathbb{P}^f}{\partial \psi_i} \mathbb{P}_{k|f} + \mathbb{P}^f \frac{\partial \mathbb{P}_{k|f}}{\partial \psi_i} \\
\frac{\partial \mathbb{P}_k}{\partial \psi_i} \frac{\psi_i}{\mathbb{P}_k} \frac{\mathbb{P}_k}{\psi_i} &= \sum_f \frac{\partial \ln(\mathbb{P}^f)}{\partial \psi_i} \mathbb{P}_{k|f} \mathbb{P}^f + \sum_f \frac{\partial \ln(\mathbb{P}_{k|f})}{\partial \psi_i} \mathbb{P}_{k|f} \mathbb{P}^f \\
\frac{\partial \ln(\mathbb{P}_k)}{\partial \ln(\psi_i)} &= \sum_f \left[\frac{\partial \ln(\mathbb{P}^f)}{\partial \ln(\psi_i)} + \frac{\partial \ln(\mathbb{P}_{k|f})}{\partial \ln(\psi_i)} \right] \frac{\mathbb{P}_{k|f} \mathbb{P}^f}{\mathbb{P}_k}
\end{aligned}$$

Analogously, for the own price elasticity we have that:

$$\frac{\partial \ln(\mathbb{P}_k)}{\partial \ln(\psi_k)} = \sum_f \left[\frac{\partial \ln(\mathbb{P}^f)}{\partial \ln(\psi_k)} + \frac{\partial \ln(\mathbb{P}_{k|f})}{\partial \ln(\psi_k)} \right] \frac{\mathbb{P}_{k|f} \mathbb{P}^f}{\mathbb{P}_k}$$

Now note that:

$$\begin{aligned}\frac{\partial \mathbb{P}_{k|f}}{\partial \psi_i} &= (\omega_{fk} T_k \psi_k^\theta)^{\frac{1}{1-\rho_f}} \left(\sum_{k'} (\omega_{fk'} T_{k'} \psi_{k'}^\theta)^{\frac{1}{1-\rho_f}}\right)^{-2} \times -\frac{\theta}{1-\rho_f} (\omega_{fi} T_i \psi_i^\theta)^{\frac{1}{1-\rho_f}} \frac{1}{\psi_i} \\ &= -\frac{\theta}{1-\rho_f} \mathbb{P}_{i|f} \frac{\mathbb{P}_{k|f}}{\psi_i} \\ \frac{\partial \ln(\mathbb{P}_{k|f})}{\partial \ln(\psi_i)} &= -\frac{\theta}{1-\rho_f} \mathbb{P}_{i|f}\end{aligned}$$

Further note that:

$$\begin{aligned}\frac{\partial \mathbb{P}^f}{\partial \psi_i} &= \frac{1}{\sum_{f'} \left[\sum_{k'} (\omega_{f'k'} T_{k'} \psi_{k'})^{\frac{1}{1-\rho_f}}\right]^{1-\rho_f}} \left[\frac{\theta}{\psi_i} \left(\sum_{k'} (\omega_{fk'} T_{k'} \psi_{k'}^\theta)^{\frac{1}{1-\rho_f}}\right)^{1-\rho_f} \frac{(\omega_{fi} T_i \psi_i^\theta)^{\frac{1}{1-\rho_f}}}{\sum_{k'} (\omega_{fk'} T_k \psi_k^\theta)^{\frac{1}{1-\rho_f}}} - \Gamma\right] \\ &= \frac{\theta}{\psi_i} \mathbb{P}^f \mathbb{P}_{i|f} - \frac{\left[\sum_{k'} (\omega_{fk'} T_{k'} \psi_{k'}^\theta)^{\frac{1}{1-\rho_f}}\right]^{1-\rho_f}}{\sum_{f'} \left[\sum_{k'} (\omega_{f'k'} T_{k'} \psi_{k'})^{\frac{1}{1-\rho_f}}\right]^{1-\rho_f}} \Gamma\end{aligned}$$

$$\begin{aligned}\Gamma &= \sum_{f'} \frac{\theta}{\psi_i} \mathbb{P}^{f'} \mathbb{P}_{i|f'} \\ &= \frac{\theta}{\psi_i} \mathbb{P}_i\end{aligned}$$

$$\begin{aligned}\frac{\partial \mathbb{P}^f}{\partial \psi_i} &= \frac{\theta}{\psi_i} \left[\mathbb{P}^f \mathbb{P}_{i|f} - \mathbb{P}^f \mathbb{P}_i\right] \\ \frac{\partial \ln(\mathbb{P}^f)}{\partial \ln(\psi_i)} &= \theta \left[\mathbb{P}_{i|f} - \mathbb{P}_i\right]\end{aligned}$$

Substituting in, we now have the following expression for the cross price elasticities:

$$
\begin{aligned}
\frac{\partial \ln(\mathbb{P}_k)}{\partial \ln(\psi_i)} &= \sum_f \left[\frac{\partial \ln(\mathbb{P}^f)}{\partial \ln(\psi_i)} + \frac{\partial \ln(\mathbb{P}_{k|f})}{\partial \ln(\psi_i)} \right] \frac{\mathbb{P}_{k|f}\mathbb{P}^f}{\mathbb{P}_k} \\
&= \theta \sum_f \left[\mathbb{P}_{i|f} - \mathbb{P}_i - \frac{1}{1-\rho_f}\mathbb{P}_{i|f} \right] \frac{\mathbb{P}_{k|f}\mathbb{P}^f}{\mathbb{P}_k} \\
&= \theta \sum_f \left[\frac{\rho_f}{\rho_f - 1}\mathbb{P}_{i|f} - \mathbb{P}_i \right] \frac{\mathbb{P}_{k|f}\mathbb{P}^f}{\mathbb{P}_k} \\
&= \theta \left[\sum_f \frac{\rho_f}{\rho_f - 1}\mathbb{P}_{i|f}\mathbb{P}_{k|f}\frac{\mathbb{P}^f}{\mathbb{P}_k} - \sum_f \frac{\mathbb{P}_i}{\mathbb{P}_k}\mathbb{P}_{k|f}\mathbb{P}^f \right] \\
&= \theta \left[\sum_f \frac{\rho_f}{\rho_f - 1}\mathbb{P}_{i|f}\mathbb{P}_{k|f}\frac{\mathbb{P}^f}{\mathbb{P}_k} - \frac{\mathbb{P}_i}{\mathbb{P}_k}\sum_f \mathbb{P}_{k|f}\mathbb{P}^f \right] \\
&= \theta \left[\sum_f \frac{\rho_f}{\rho_f - 1}\mathbb{P}_{i|f}\mathbb{P}_{k|f}\frac{\mathbb{P}^f}{\mathbb{P}_k} - \frac{\mathbb{P}_i}{\mathbb{P}_k}\mathbb{P}_k \right] \\
&= \theta \left[\sum_f \left(\frac{\rho_f}{\rho_f - 1}\mathbb{P}_{i|f}\mathbb{P}_{k|f}\frac{\mathbb{P}^f}{\mathbb{P}_k} \right) - \mathbb{P}_i \right] \\
&= \frac{\theta}{\mathbb{P}_k} \left[\sum_f \left(\frac{\rho_f}{\rho_f - 1}\mathbb{P}_{i|f}\mathbb{P}_{k|f}\mathbb{P}^f \right) - \mathbb{P}_i\mathbb{P}_k \right]
\end{aligned}
$$

Note that since $\rho_f \in [0, 1)$, it must be that: $\frac{\rho_f}{1-\rho_f} < 0$. Since $\mathbb{P}_i \in (0, 1)$ and $\theta > 2$, this implies that $\frac{\partial \ln(\mathbb{P}_k)}{\partial \ln(\psi_i)} < 0$. This is essentially the gross-substitutes property and is crucial for proving the uniqueness of the equilibrium.

To derive the own price elasticity note that:

$$
\begin{aligned}
\frac{\partial \mathbb{P}_{k|f}}{\partial \psi_k} &= \frac{1}{\sum_{k'} (\omega_{fk'} T_{k'} \psi_{k'}^\theta)^{\frac{1}{1-\rho_f}}} \left[(\omega_{fk} T_k \psi_k^\theta)^{\frac{1}{1-\rho_f}} \frac{\theta}{\psi_k} - \mathbb{P}_{k|f} (\omega_{fk} T_k \psi_k^\theta)^{\frac{1}{1-\rho_f}} \frac{\theta}{\psi_k} \right] \\
&= \theta \frac{\mathbb{P}_{k|f}}{\psi_k} \left[1 - \mathbb{P}_{k|f} \right] \\
\frac{\partial \ln(\mathbb{P}_{k|f})}{\partial \ln(\psi_k)} &= \theta(1 - \mathbb{P}_{k|f})
\end{aligned}
$$

$$\begin{aligned}
\frac{\partial \ln(\mathbb{P}_k)}{\partial \ln(\psi_k)} &= \sum_f \left[\frac{\partial \ln(\mathbb{P}^f)}{\partial \ln(\psi_k)} + \frac{\partial \ln(\mathbb{P}_{k|f})}{\partial \ln(\psi_k)}\right] \frac{\mathbb{P}_{k|f}\mathbb{P}^f}{\mathbb{P}_k} \\
&= \sum_f \left[\theta(\mathbb{P}_{k|f} - \mathbb{P}_k) + \theta(1 - \mathbb{P}_{k|f})\right] \frac{\mathbb{P}_{k|f}\mathbb{P}^f}{\mathbb{P}_k} \\
&= \sum_f \theta\left(1 - \mathbb{P}_k\right) \frac{\mathbb{P}_{k|f}\mathbb{P}^f}{\mathbb{P}_k} \\
&= \theta \frac{1 - \mathbb{P}_k}{\mathbb{P}_k} \sum_f \mathbb{P}_{k|f}\mathbb{P}^f \\
&= \theta(1 - \mathbb{P}_k)
\end{aligned}$$

Since $\mathbb{P}_k \in (0, 1)$, this is strictly bounded below by 0 and above θ.

E.2 Additional Plots

F Computational Approach

This is a large system of equations, and solving the system requires a bit of numerical elbow grease. This section describes the general strategy I use to solve the model. The first step is to derive the appropriate discretization strategy for the differential equations. In general, it is important to respect the direction of a differential equation when solving it. Forward ODEs (those with boundary conditions at time T) should be integrated backwards, while backwards ODEs (those with boundary conditions at time 0) should be integrated forward. Writing the discretized ODEs in the appropriate way generates the following set of equations:

$$\tilde{\mathbf{L}}_{kt} - \tilde{\mathbf{L}}_{k,t-\Delta t} = \delta\big(\mathcal{E}_{kt} - \tilde{\mathbf{L}}_{kt}\big)\Delta t$$
$$\tilde{\mathbf{L}}_{kt}(1 + \delta\Delta t) = \tilde{\mathbf{L}}_{k,t-\Delta t} + \delta\mathcal{E}_{kt}\Delta t$$

$$\tilde{\mathbf{L}}_{kt}(1 + (\delta + \lambda_k)\Delta t) = \tilde{\mathbf{L}}_{k,t-\Delta t} + \delta\mathcal{E}_{kt}\Delta t$$

And for the forward equations, we have:

$$\tilde{\mathbf{H}}_{kt}(1 + (r_t + \delta - g_t)\Delta t) = \tilde{\mathbf{H}}_{\mathbf{k,t+\Delta t}} + \Delta t\ w_{kt}$$

$$\hat{\mathbf{H}}_{kt}(1 + (r_t + \delta + \lambda_k - g_t)\Delta t) = \hat{\mathbf{H}}_{\mathbf{k,t+\Delta t}} + \Delta t\ w_{kt}$$

$$\mathcal{V}_{st}(1 + (r_t + \Lambda_{st} - g_t)\Delta t) = \mathcal{V}_{s,t+\Delta t} + (\pi_{st} - \epsilon_{st}\tilde{w}_{st}\Lambda_{st}^{\gamma})\Delta t$$

At any given point in time, we need to solve for the within period competitive equilibrium. This requires solving for ι_{st}, w_{kt}, $\mathbf{Y}_t$,r_t, and g_t using the following system of equations:

$$\mathbf{Y}_t = AL_t^P$$

$$\mathbf{C}_t = \mathbf{Y_t}$$

$$\mathbf{C}_t = (\rho + \delta)(\mathbf{A}_t + \sum_k \mathbf{H}_{kt})$$

$$g_t = \ln(\lambda) \sum_s \beta_s \iota_s$$

$$\hat{w}_{kt}\mathbf{L}_{kt} = \frac{1}{\gamma} \sum_s \alpha_{sk} V_{st} \iota_{st}$$

$$\iota_s = \left[\frac{\gamma \epsilon_{st} \tilde{w}_{st}}{V_{st}} \right]^{\frac{1}{1-\gamma}}$$

$$\mathcal{E}_{kt} = \mathbb{P}_{kt} \mathbb{E}_t[z_k | \text{choose } k, \tau = 0]$$

$$\mathbb{P}_{kt} = \sum_f \left\{ \frac{\left[\sum_{k'} (\omega_{fk'} \psi_{k'}^{\theta})^{\frac{1}{1-\rho_f}} \right]^{1-\rho_f}}{\sum_f \left[\sum_{k'} (\omega_{fk'} \psi_{k'}^{\theta})^{\frac{1}{1-\rho_f}} \right]^{1-\rho_f}} \times \frac{(\omega_{fk} \psi_k^{\theta})^{\frac{1}{1-\rho_f}}}{\sum_{k'} (\omega_{fk'} \psi_{k'}^{\theta})^{\frac{1}{1-\rho_f}}} \right\}$$

Notice that these equations depend only on time t state variables, and are solvable in such.[15] This suggests the standard approach: at each time t step, one solves the set of static equilibria equations to recover the terms that enter into the dynamic equations. Solving for the aggregate paths of such states requires perturbations to the static equilibra at each point in time, which can then be computed efficiently via the implicit function theorem.

Though the algorithm is straightforward, the dimensionality of the problem requires additional

[15] Formally, this system is an index-1 differential algebraic system of equations.

numerical tricks. Specifically, one can leverage the fact that because the dynamic linkages are either one period forward or backward, the Jacobian of this system will have a particular block tridiagonal form. This linkage is valuable for two reasons. Firstly, computing these matrices can be done quickly through automatic differentiation. Secondly, once the Jacobian has been computed, one can very efficiently invert this system by making use of the (block-)Thomas algorithm. With the Jacobian inverted, one can then perform a standard geometric line-search at each Newton step.[16]

To add taxes to the system requires solving for the labor tax τ_{0t} that makes the government budget constraint hold with equality at each point in time:

$$\sum_k \tau_{kt} \hat{w}_{kt} \mathbf{L}_{kt} = 0$$

[16]In theory, the Thomas algorithm can be unstable, although if the matrix to be inverted has a dominant diagonal or is positive semi-definite then the algorithm is stable. In practice, I have not run into problems with stability, consistent with the general commentary in Press (2007).

G Data

To be done once statistics are disclosed.